



PAID FOR  
NOT TO BE USED

CHECKED-75

Acc. No.	19381
Class No.	D.1.
Book No.	43

1871183

Checked





ASAFTYAH STATE LIBRARY

Hyderabad-Deccan

Author .....

Title .....

Acc. No. ....

Class No. ....

Book No. ....

301

DOD

Call No. ....

Author *DOD S.*

Title *Dimensions  
of Society*

RARE BOOKS  
NOT TO BE LOANED

CHECKED-75

Acc. No.	19381
Class No.	D.1.
Book No.	43

67/181/83

Checked  
1987.









Dimensions



of Society





THE MACMILLAN COMPANY  
NEW YORK • BOSTON • CHICAGO • DALLAS  
ATLANTA • SAN FRANCISCO

MACMILLAN AND CO., LIMITED  
LONDON • BOMBAY • CALCUTTA • MADRAS  
MELBOURNE

THE MACMILLAN COMPANY  
OF CANADA, LIMITED  
TORONTO

## SUMMARY OF FORMULAE OF S-THEORY

**General Formula:** The generalization, "People's characteristics and environments change," can be more rigorously stated as: "Any quantitatively recorded societal situation (S) can be expressed as a combination of:

4 indices ((I)), namely: of time [T], space [L], a human population [P], and indicators [I] of their characteristics; each modified by

4 scripts, namely: the exponent [<sup>s</sup>], and descripts denoting a series of classes [<sub>s</sub>], of class-intervals [<sub>s</sub>], and of cases [<sup>s</sup>]; all combined by

8 operators (;), i.e.: for adding [+], subtracting [-], multiplying [×], dividing [/], aggregating [:], cross-classifying [::], correlating [·], and identifying [']."

The S-theory is a system of hypotheses which assert that combinations of these basic concepts [in square brackets] will describe and classify every tabulation, graph, map, formula, prose paragraph, or other set of quantitative data in any of the social sciences.

### Sectoral Formula:

(expanding the general formula by sectors)

cases = s      s = exponents  
(I') = homosectoral index  
class-intervals = s      s = classes

$${}_s(I')_s^s = \left\{ \begin{array}{ll} \begin{array}{l} \text{dates} = t \quad t = \text{temporal} \\ \quad \quad \quad \text{exponent} \\ \quad \quad \quad \text{or} \\ \quad \quad \quad T = \text{temporal} \\ \quad \quad \quad \text{units} \end{array} & \begin{array}{l} \text{points} = l \quad l = \text{spatial ex-} \\ \quad \quad \quad \text{ponent} \\ \quad \quad \quad \text{or} \\ \quad \quad \quad L = \text{spatial units} \end{array} \\ \begin{array}{l} \text{periods} = t \quad t = \text{chronometries} \\ \quad \quad \quad \text{or} \\ \text{points} = i \quad i = \text{indicatory} \\ \quad \quad \quad \text{exponent} \\ \quad \quad \quad \text{or} \\ \text{class-} \quad \quad \quad I = \text{indicatory} \\ \text{intervals} = i \quad i = \text{indicators} \end{array} & \begin{array}{l} \text{sub-spaces} = l \quad l = \text{spaces} \\ \quad \quad \quad \text{or} \\ \text{persons} = p \quad p = \text{populational} \\ \quad \quad \quad \text{exponent} \\ \quad \quad \quad \text{or} \\ \text{class-} \quad \quad \quad P = \text{persons} \\ \text{intervals} = p \quad p = \text{plurels} \end{array} \end{array} \right.$$

**Quantic Number:** (expanding the general formula by exponents only)

$$|s| = |t; i; l; p|$$

### Societal Situations and Societal Space:

In symbolic terms—      S      :      (I')      :      |<sub>s</sub>      :      |<sup>s</sup>      :      |<sub>s</sub>      :      |<sup>s</sup>

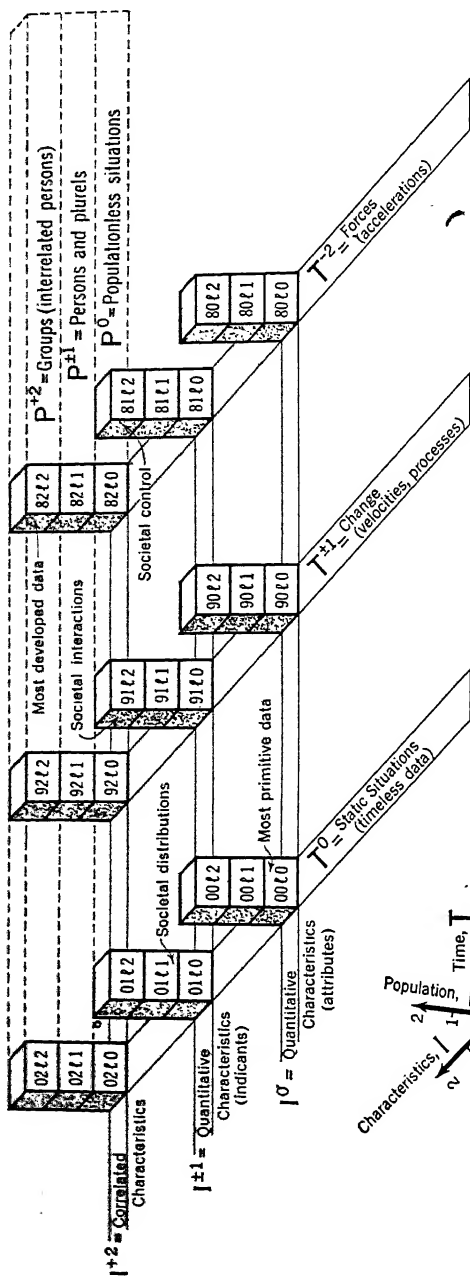
In geometric terms—      Any portion of S-space      :      sectors      :      vectors      :      normals      :      sects      :      points

In sociologic terms—      Any quantitative societal situation is analyzable into 4 types of indices T, I, L, P each type analyzable into a specified number of indices each operationally developed by its exponents and each subdivided into a specified number of class-intervals and further subdivided into a specified number of cases

### Some Standard Applications: (by Chapters 3-12)

	Quantic number =		<sup>s</sup> =
I <sub>0</sub> = a series of qualitative characteristics	0;0;0;0	P <sub>0;q;r</sub> = a classification of plurels	0;0;0;1
iI:P = a frequency distribution of persons	0;1;0;1	σI, .σI,, = correlated characteristics	0;2;0;0
°P:°P:I = a group, interrelated persons	0;1;0;2	PL <sub>1</sub> <sup>2</sup> = ecological densities	0;0;8;1
tT:P = a population pyramid	1;0;0;1	(I)T <sup>-1</sup> = a societal process	9;1;0;0
T <sup>-2</sup> IP = a societal force	8;1;0;1	°P:°P:IT <sup>-2</sup> = societal control	8;1;0;2

## THE "QUANTIC SOLID" OF S-THEORY



Each block represents one class of quantitatively expressed societal situations. The "quantic number" identifying each block is composed of the exponents on the four indices in the quantic formula:  $S = T^t i^i l^l p^p$

The four sectors of societal situations, S

CHICAGO - 1968

# Dimensions of Society

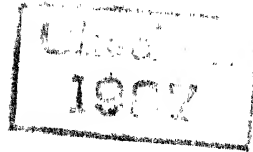
A QUANTITATIVE SYSTEMATICS  
FOR THE SOCIAL SCIENCES

*Not to be issued*

by Stuart Carter Dodd

Chairman of the Department of Sociology  
American University of Beirut

DATE BOOK  
NOT TO BE ISSUED



NEW YORK • 1942

THE MACMILLAN COMPANY

COPYRIGHT, 1942  
BY THE MACMILLAN COMPANY

*All rights reserved—no part of this book may be reproduced in any form without permission in writing from the publisher, except by a reviewer who wishes to quote brief passages in connection with a review written for inclusion in magazine or newspaper*

PUBLISHED JANUARY, 1942

Acc. No.	19381
Class No.	D.1.
Book No.	43

الى الزايت

انما وان لم تكن خيرة في علم الاجتماع  
الرياضي فقد ساعدت بتأثيرها على المناشئ

## PREFACE

### *Scope*

This study of the theory which is specified by the formula  $S = \{ (T; I; L; P) \}$ , though emphasizing the field of Sociology as currently conceived, is applicable to all quantifiable data in each of the social sciences. It is a mathematical approach to society; an attempt to systematize statistical forms and societal data.

It is intended for the rising generation of social scientists who are demanding tools of increased objectivity and precision for dealing with the phenomena of society.

This volume is a companion volume to *Foundations of Sociology* by George A. Lundberg. The postulates and methodological principles which are there more ably and more fully developed are the foundation for this study of the dimensions of society.

### *Presentation*

The facts from which this theory was induced are represented by the 326 "S-situations" which are appended at the ends of the chapters. Each "S-situation" is a unitary set of quantitative societal data (referred to in a serial number within each chapter by the abbreviation "S. 1, Ch. 2," etc.).

In Chapter 2 the theory generalizing and classifying these situations is stated; its possible utility in the social sciences is discussed; and the experiments made towards verifying it are described.

The next eleven chapters explore this classificatory theory systematically, class by class, developing operational redefinitions of sociological concepts and relations between them in some seven hundred formulae. (These formulae are numbered consecutively as equations within each chapter and referred to in abbreviated form as "Eq. 1, Ch. 1," etc.) In order to reduce the difficulty which most social scientists have, of learning an algebraic symbolism, the more detailed and technical formulae with the discussion of them have been relegated to footnotes (referred to as "Note 2, Ch. 3" etc.) and placed at the ends of chapters. A glos-

## PREFACE

vii

sary of symbols appears in the appendices, where it is available to the advanced student and will interrupt the attention of the general reader less. The appendices include a bibliography to which the reader is referred by a note, "Ref. 1," etc. in the text.

The last chapter attempts a tentative appraisal of the theory as a contribution towards making sociology more of an exact science.

### *An acknowledgment*

The only acknowledgment which is more than the usual courtesy, is to the author's tutor in sociological research. This tutelage began by studying the book *Social Research* by George A. Lundberg and by applying its teaching in a controlled experiment on village communities in Syria. The tutelage in sociological methodology continued on meeting George Lundberg later in the States and during a subsequent visit together one summer in the Lebanon mountains. His interest, advice, and encouragement have been the chief stimuli in making this study.

STUART C. DODD

*Beirut, Syria*  
*August, 1941*

## PREFACE

### *Scope*

This study of the theory which is specified by the formula  $S = (T; I; L; P)_s$ , though emphasizing the field of Sociology as currently conceived, is applicable to all quantifiable data in each of the social sciences. It is a mathematical approach to society; an attempt to systematize statistical forms and societal data.

It is intended for the rising generation of social scientists who are demanding tools of increased objectivity and precision for dealing with the phenomena of society.

This volume is a companion volume to *Foundations of Sociology* by George A. Lundberg. The postulates and methodological principles which are there more ably and more fully developed are the foundation for this study of the dimensions of society.

### *Presentation*

The facts from which this theory was induced are represented by the 326 "S-situations" which are appended at the ends of the chapters. Each "S-situation" is a unitary set of quantitative societal data (referred to in a serial number within each chapter by the abbreviation "S. 1, Ch. 2," etc.).

In Chapter 2 the theory generalizing and classifying these situations is stated; its possible utility in the social sciences is discussed; and the experiments made towards verifying it are described.

The next eleven chapters explore this classificatory theory systematically, class by class, developing operational redefinitions of sociological concepts and relations between them in some seven hundred formulae. (These formulae are numbered consecutively as equations within each chapter and referred to in abbreviated form as "Eq. 1, Ch. 1," etc.) In order to reduce the difficulty which most social scientists have, of learning an algebraic symbolism, the more detailed and technical formulae with the discussion of them have been relegated to footnotes (referred to as "Note 2, Ch. 3" etc.) and placed at the ends of chapters. A glos-

## PREFACE

vii

sary of symbols appears in the appendices, where it is available to the advanced student and will interrupt the attention of the general reader less. The appendices include a bibliography to which the reader is referred by a note, "Ref. 1," etc. in the text.

The last chapter attempts a tentative appraisal of the theory as a contribution towards making sociology more of an exact science.

### *An acknowledgment*

The only acknowledgment which is more than the usual courtesy, is to the author's tutor in sociological research. This tutelage began by studying the book *Social Research* by George A. Lundberg and by applying its teaching in a controlled experiment on village communities in Syria. The tutelage in sociological methodology continued on meeting George Lundberg later in the States and during a subsequent visit together one summer in the Lebanon mountains. His interest, advice, and encouragement have been the chief stimuli in making this study.

STUART C. DODD

*Beirut, Syria*  
*August, 1941*

# TABLE OF CONTENTS

SUMMARY OF FORMULAE OF S-THEORY,  $S = \{T_t^t; I_t^t; L_t^t; P_t^p\}$

*Frontispiece*

PAGE

PREFACE

vi

## PART I. THE SECTORS OF SOCIETY, (I')

### CHAPTER

#### I. SCIENCE AND SOCIOLOGY

3

an introductory specification of the purpose of this volume as a contribution to a quantitative, systematic science of Sociology.

#### II. THE S-THEORY, S

25

a classification of societal situations by means of S-theory; its possible utility; and its verification by three technics of testing internal consistency and by testing its correspondence to observed data; with 35 illustrative situations reduced to S-formulae.

## PART II. THE CHARACTERISTICS SECTOR, $I_t^t$

#### III. INDICATORS, $I^t$

135

classification of the characteristics of people and their environment by completeness of observation (qualitative, quantitative, correlative) as symbolized by the exponent; by content (physical, biological, cultural, and subclasses) as symbolized by the class-script; by precision (all-or-none, ordinal, cardinal) as symbolized by the class-interval and case scripts; together with the geometric interpretation of these scripts; a sample of situations used to induce the formulae of this chapter.

## PART III. THE POPULATION SECTOR, $\frac{P_t^p}{P_t^p}$

#### IV. PLURELS, P

203

classification of population by qualitative characteristics as symbolized by the exponent; the class or plurel script, the class-interval script, and the case or person script, together with their geometric interpretation; a sample of situations used to induce the formulae of this chapter.

#### V. DISTRIBUTIONS, $I_t^p$

242

classification of population by quantitative characteristics and by quantitatively evaluated characteristics; "social pathology" as extreme negative valuations; a tension theory of societal action in respect to values; a sample of situations used to induce the formulae of this chapter.

#### VI. CORRELATIONS, $I^t$

307

correlation as a product of characteristics (of people or of their environment); probability and causation as forms of correlation; correla-

# TABLE OF CONTENTS

ix

## CHAPTER

## PAGE

tions resolved into components (or factors) and their elements; 24 systematic diagrams interpreting correlation; a sample of situations used to induce the formulae of this chapter.

### VII. INTERRELATIONS, $P^2$

389

a group as interrelated people; interrelations classified by sectors, by the form of the surface of the matrix, by internal structure of the group; a sample of situations used to induce these formulae.

### *PART IV. THE SPACE SECTOR, $\{L\}$*

### VIII. DENSITIES, $L^1$

449

a societal classification of geographic spaces as symbolized by the four spatial scripts; societal densities as defining rural and urban sociology and ecological concepts; a sample of situations used to induce these formulae.

### *PART V. THE TIME SECTOR, $\{T\}$*

### IX. DURATIONS, $T^{+1}$

495

a temporal classification of societal situations as symbolized by the four temporal scripts; the duration of populations and of their indicators; Brief S formulae; a sample of situations used to induce the formulae of this chapter.

### X. CHANGE, $T^{-1}$

528

societal processes, simple, complex, and aggregated, subclassified by sectors, exponents, class scripts, and signs; a sample of situations used to induce the formulae of this chapter.

### XI. FORCES, $T^{-2}$

736

accelerations and decelerations of change defining societal forces and societal control; a sample of situations used to induce the formulae of this chapter.

### *PART VI. THE SYNTHESIS OF SECTORS, $S$*

### XII. APPRAISAL OF S-THEORY, $S_v$

831

the contribution of S-theory towards understanding societal phenomena in terms of the objectivity, precision, parsimony, and comprehensiveness of the S-system of concepts with its present limitations; the contribution of S-theory towards predicting and controlling societal phenomena; with illustrative situations reduced to S-formulae.

## APPENDICES

### I. Glossary of Symbols of S-theory

899

### II. Rules for Writing S-formulae

903

### III. Research Suggestions

916

### IV. Bibliography

923

## INDEX OF TOPICS AND NAMES

929



*PART I*

THE SECTORS OF SOCIETY, (I')



## Chapter I

### SCIENCE AND SOCIOLOGY

#### I. PURPOSE OF THIS VOLUME

Our working hypothesis is, that it is possible with our present knowledge to begin constructing a *quantitative systematic science* of *sociology*. This volume presents evidence towards verifying this hypothesis. To the extent that readers judge the evidence to be adequate, the hypothesis will be positively verified; to the extent that the evidence is judged to be unsatisfactory, the hypothesis will be disproven or at least remain unproven.

The postulates underlying this hypothesis have been ably developed in Professor George A. Lundberg's companion volume, *Foundations of Sociology*. We proceed to build on these scientific points of view in offering definite technics towards further developing Sociology as a science. As the fundamental issues have been discussed by him, we shall try merely to state tersely in this chapter, the purpose of this volume, without further discussion of the philosophical and scientific issues involved. To make our purpose clear, our notion of the four terms "science," "sociology," "quantitative," and "systematic" will be briefly stated before proceeding to present a theory, or system of classificatory concepts and notation, towards accomplishing this purpose.

#### II. THE CONCEPT OF SCIENCE (IN THIS VOLUME)

##### A. The Function of Science

Science is seen as a means of human adjustment. It is a body of technics, ranging from the very particular to the most generalized, by which man can make those adjustments in himself or in his environment which he considers to be more satisfactory. This continual process of adjusting can be more analytically described as composed of (a) understanding, (b) predicting, and (c) controlling phenomena. Insofar as we understand phenomena, can predict their future occurrence, and can control them as we

desire, we are able to adjust well in our environment. This triple function of science begins with increased understanding, proceeds to predict best that which we understand most, and goes on, as far as possible, towards changing such phenomena as we desire.

In order to test the S-theory presented in this volume by these three functions of science each function will be analyzed into its constituents. Thus our understanding of phenomena depends in large part upon the *concepts* whereby we deal with phenomena and upon the *system*, weaving those concepts together. The concepts must be operationally defined, as far as possible, in order that their degree of precision-to-ambiguity and their degree of mutual-exclusion-to-overlap in meaning may be experimentally measured and better concepts progressively developed. The objectivity, the parsimony, the precision, and the comprehensiveness of the sociological concepts, almost two hundred in number, which S-theory thus defines towards improving our understanding of societal phenomena will be reviewed in Chapter XII.

Here the reader is notified that these three functions of science—to understand, predict, and control—will be the standard by which this, or any other, theory should have its scientific utility appraised. Chapters II–XI inclusive present evidence. Chapter XII summarizes this evidence to enable the student to appraise the utility of the S-theory for science.

### *B. The Content of Science*

Since the possible adjustments in the universe seem innumerable to man's limited span of attention and memory, it is expedient to classify them, dividing them up into related subfields which we call the several sciences (see S. 5, Ch. II). By specializing thus, we adjust to the vastness of the total problem of adjustment.

But science is not the only means of adjustment in our environment, for philosophy, religion, "common sense," custom, and other ways, partly overlapping, partly diverse, are also available. Science then is only one means of adjusting. What specifies this means seems to us its function, stated above, together with the process of building a science. This process is usually called the scientific method. A science is the knowledge derived by the scientific method. The method, not the content, defines the body

of knowledge as a science. Any kind of knowledge—astronomical or astrological, chemical or alchemic, psychological or phrenological, sociological or proverbial—may *become* a science by applying the scientific method to the phenomena that are its subject matter. In these terms, “science” in this volume denotes sensory data that have undergone a process of being recorded, systematized, and verified. Each of these steps in converting sensory data into a science, requires some qualifying, lest brevity of statement breed misunderstanding.

### *C. The Scientific Method*

The general steps of the scientific method have been variously described. In primitive form, they are Dewey’s analysis of problem-solving (Ref. 11) into (1) a felt difficulty, (2) its definition, (3) suggestions for solution, (4) their implications, and (5) decision and trial of one suggestion. We shall summarize the steps in more rigorous scientific work under the four headings of Problem, Observation, Systematization, and Verification, as a generalized and somewhat over-simplified pattern from which there are many variants.

#### 1. THE PROBLEM

Scientific work begins with some problem, be it a question to be answered, an hypothesis to test, an instrument or technic to be developed, or a field of phenomena to explore. In rigorous work the problem will be carefully formulated in words or other symbols, its limits specified, its terms defined. This we attempt to do in this volume, first in stating our working hypothesis as the possibility of constructing a quantitative systematic science of Sociology, and then in expounding each of those specifications.

#### 2. OBSERVATION

With the problem or field of investigation defined, the next step is to observe all the relevant phenomena to get all the pertinent facts. These observations are sensory data, which are then interpreted and related to previous sensory data. Instruments, whether microscopes for bacteria, or schedule cards for societal trends, extend our senses, enabling us to observe phenomena to which our naked senses are insensitive or inadequate.

The operational procedure, what to do first, second, and third, is just as important a part of observing as are the material instruments. The instructions as to how to proceed in analyzing the unknown chemical substance, in finding a child's intelligence quotient, or a community's hygienic rating, are essential to the accuracy of the observations. Furthermore, the observations to be facts, must be objective, not subjective ones, i.e., they must be in a communicable form and agreed upon by different observers. An item of observation becomes a fact at a given time, in proportion to the percentage of competent observers who agree upon it, i.e., who respond similarly to that stimulus-situation. Recording facts tends to increase their objectivity, as well as preserving them in order that they can be tested by reobservations and communicated to others. Recording also cumulates the facts into a growing body of knowledge which is the basis of a science.

This step of collecting the relevant facts has been taken in the study reported in this volume in the form of collecting graphs, tables, maps, diagrams, and other quantitative sets of data from the literature of the social sciences during the past twenty years. Some fifteen hundred sets of such data were taken as a sample of quantitative facts which any systematic theory must fit and organize into some sort of unity. Three hundred of these sets of data, which we shall refer to as societal *situations*, are recorded in this volume as a sample of the facts to be systematized by some theory.

### 3. SYSTEMATIZATION

The facts are then summarized. Many particular items are generalized. Principles are induced. These generalizations may take various forms, such as in:

#### a. *A concept*

This is a name for a class of percepts having common characteristics. (A percept is always completely present to the senses at one time—a concept never is completely present, as it includes the possibility of other similar percepts. Thus, a visual percept can be recorded in a photograph, an auditory one on a gramophone disc.)

*b. A statistical index*

This is a numerical concept, a summary of many numbers in a single number defined by its computational formula.

*c. A classification*

This is an arrangement of concepts in a hierarchy, in which a concept is subdivided into classes and these into subclasses (see S. 5, Ch. II; S. 1, Ch. III; S. 2 and S. 10, Ch. IV; S. 1, Ch. IX). Every frequency distribution is a classification in which the classes are varying amounts of one concept, rather than qualitative subclasses.

*d. A statement of a generalized relation*

This may be static or functional, and either descriptive or causal. If unverified this may be an hypothesis; if positively verified and of utility for prediction and control of phenomena, it is called a scientific law, e.g., the law of gravity, Engel's law, that the proportion of family income spent on necessities tends to vary inversely with the size of income (see S. 47, Ch. X, for a crude presentation), the law of sampling, that accuracy of prediction varies with the square root of the size of the sample (Eq. 14, Ch. IX).

Obviously scientific laws, which are a brief, exact, and idealized statement in a mathematical formula of a relation that is true under specified conditions, vary from such highly exact statements as those about the planetary orbits, to probability statements, as in actuarial insurance. Both deserve to be called laws, provided the attendant conditions and degree of probability are specified sufficiently, so that predictions of the future can be made within known limits of error.

The generalized relation may be a causal principle such as that the typhoid bacillus under specified conditions will cause typhoid in a person, or that violation of a group's mores will cause action, i.e., stimulate a response, to preserve its mores under conditions of the violation being a maximal one and the particular mores being maximally valued in the absence of other restraints upon the group. ("Maximal" is here defined as greater than two standard deviations above the mean on the appropriate scale.)

Generalized relations are of many kinds and no classification of them is attempted here beyond a remark or two intended to amplify our notion of systematizing as a step in the scientific method.

*e. A theory, or system of concepts and generalized relations*

This is an orderly and internally consistent organization of all the facts and their generalizations in some limited field. Thus the S-theory, which is proposed in this volume, is an internally consistent organization of about a dozen symbols for expressing in generalized form many of the concepts of Sociology, for expressing a rigorous classification of all quantitative sociological facts by means of operationally related statistical indices, and for increasing the precision of whatever quantitative generalizations can be made with our current all-too-inadequate data.

Generalizations are sought because our human capacity for responses is limited to a far smaller number than the millions of stimuli and situations in the world about us. But proportionally as we can generalize these into classes their number is reduced nearer to our capacity and we can adjust to a larger proportion of existent and potential situations. This is the psychological reason why science seeks laws. If the adjustment capacities of people were not finite, adjusting to a billion different particular situations would be as possible as to a thousand classes of situations.

But generalizations hold only under specified conditions, and these may vary. This means that generalizations are related in patterns. Hence, to use them, we need to know those patterns and this is what a theory, a system of generalizations, aims to do. It symbolizes a set of interworking generalizations so that our understanding, prediction, and control of the phenomena covered by that theory may be more adequate.

#### 4. VERIFICATION

Generalizations related and combined into a system may be true or imaginary, and may meet the pragmatic test of enabling man to predict and control phenomena in any degree between zero and perfection. Verification is the step of determining the degree of truth, the percentage of perfect prediction and control,

given by a generalization or by a theory. How adequately does this theory work? The final step of a cycle of scientific work answers this question and, of course, thereby formulates a problem for another cycle of scientific work, in trying to improve the adequacy of that theory or of a substitute one.

There are definite technics of verification in great variety in every science. The fundamental method, to which more specific technics contribute, is to try predicting and controlling phenomena by deductively applying the generalization (or system of them) to particular situations. If these bacilli cause typhoid, injecting them into this person will give him typhoid. Inject them and see if the predicted typhoid results. If maximal violation of maximally valued mores stimulates a group to respond by eliminating the violator with a specified probability, then, in the next hundred violations observed, elimination of the violator must follow in the percentage given by that specified probability.

Whenever the conditions, i.e., characteristics of a situation, can be varied separately by human effort, a crucial experiment to verify a theory is possible. Whenever the conditions vary but cannot be fully varied, i.e., controlled by human effort, statistical selection may be possible in order to isolate characteristics or sets of them, and so test a theory. In either type, together with intermediate variants, a deductive prediction is made from previous generalizations as to what further observations should be, whether of the manipulated or of the selected sort. The degree of agreement between these predicted observations and subsequent actual observation of the situations as they transpire is the climax of the verifying process.

The verification, however, includes more than a final testing of generalizations. It includes as corollaries, technics which verify the previous steps and contribute to verifying the final result. Thus, the observations must be verified by repeating them with different observers and instruments, at differing times and places, and under other differing conditions. Determining the reliability and limits of the observations is part of the verifying process. It establishes the facts.

Under the systematizing step, verifying may include determining the objectivity of the concepts and of the classification in the hands of different scientists. Thus a pioneer, controlled

experiment to determine the objectivity of the concepts and classification provided by a sociological theory is reported in Chapter II.

## 5. OPERATIONAL DEFINITIONS

An important technic towards verifying is to give definitions in operational terms. Sociologists are still wedded to definitions in descriptive terms, which do not enable another party to duplicate the thing defined. A definition which tells what to do first, second, third, with specified materials, in order to get the thing defined, is for science as far superior to a mere description of its properties as a kitchen recipe for making a cake is superior, *if one wants to get a cake*, to a literary eulogy of its color, texture, and delicious taste. Thus, "competition" may be described and evaluated by adjectives such as "selfish," "cutthroat," "capitalistic," "individualistic struggle," etc., but we define it operationally as the process measured by calculating the standard deviation of percentage gains and losses of a desideratum, V, competed for among the competitors, P, in a period, T (see Eq. 47a, Ch. X). This is our model for defining all the two hundred odd terms for Sociology and for the social sciences generally in this volume. Such definitions enable another investigator to verify more exactly whether the phenomenon that is defined exists as asserted.

Reliance on operational definitions has the result of so radically rearranging the traditional content of a textbook on Sociology, as to lead to dismay among many professional sociologists. Such readers will look over the graphs in this volume, for example. These graphs are classified by the operation of the exponent into the quantic classification, the degree of perfection of which will be expounded and experimentally measured in Chapter II. But as the graphs will not be grouped around familiar rubrics of "delinquency," "the family," "rural communities," "culture," and the like, they will seem to those readers to be grouped very heterogeneously. Such readers are warned that operational principles are the first demand of pure science, their application to human problems is technology and should be reserved for a later stage. We deal in this volume with human characteristics first, then with the operational combination of these with people, de-

fining plurels and distributions, correlations of characteristics in patterns, and interrelations of people, each as defined by an operational formula. Then spatial and temporal principles in ecological and dynamic situations are studied as specified by their calculational formulae of densities, velocities, and accelerations of change. Finally any combination of these principles can be applied appropriately to any given societal problem or field.

This is what every mature science does. Physics textbooks do not study "thunderstorms" all in one place. The principles of sound waves in thunder will be studied where sound is studied with operational technics for such phenomena whether manifested in thunder or in an oratorio of the Messiah. The principles of electric charges are elsewhere grouped and the principles of light in another section, regardless of whether it occurs in lightning or in a searchlight. These principles may be combined to explain the thunderstorm as a whole, or they may be combined by the engineer to build a subway. But the principles are best studied as grouped by similarity of operation, not by content of application. Sociologists are still too much interested in social work, in *immediate* application of theory to societal situation, to realize fully that a longer perspective requires putting aside applications and first building up a body of several hundred concepts and principles which are operationally defined by their statistical form (i.e., formulae) and general to any societal content. This is what this volume attempts to do. But unless the reader grasps this operational point of view, he will misjudge the societal principles and their organization throughout the rest of this volume.

Verifying of the systematizing step may further include checking statistical calculations, criticizing the formulae used, or the logic leading to the inductions or to the deductive applications. All this tends towards establishing the generalizations as truth, subject to the more final test of accurate prediction and adequate control of the phenomena generalized.

The above sketch of our notion of scientific method is an oversimplification of a typical pattern. There are many variants in scientific work on particular problems. Years of work may be concentrated on one step as in perfecting an instrument or a technic of observing. A generation of professional discussion may be largely wasted on a sterile formulation of a problem, as in the

search to identify "instincts" in human Psychology. The sketch over-simplifies in that the four steps interwork and do not always follow the sequence as described above. A problem may, and usually does, become better formulated as observations and tentative systematizing go forward. The systematizing redirects the observer to collect further facts with better technics under other conditions, etc. Each of the steps may fertilize and influence the others in complex ways. Their intellectual isolation for discussion, however, is useful to make each step more definitely crystallized, communicable, and open to critical improvement and evaluation by fellow scientists. As an exposition of scientific method, this sketch is far too brief. However, the many implications and issues involved in its somewhat dogmatic statements have been more fully dealt with in Lundberg's companion volume, *Foundations of Sociology*. Therefore, a mere outline is presented here to suggest the premises upon which the theory is based in the succeeding chapters.

### III. THE CONCEPT OF SOCIOLOGY (IN THIS VOLUME)

Sociology as the science of society, where, by "society" is meant the whole of human life, needs further restricting else Sociology would be simply the total of the social sciences. In practice, the content of Sociology is operationally defined by the studies published by sociologists, such as the members of the national sociological associations. For a conceptual summary of this field Sorokin's definition (Ref. 66, p. 760) usefully draws the boundaries around three mutually complementary fields of knowledge.

"It seems to be a study, first, of the relationship and correlations between various classes of social phenomena, (correlations between economic and religious; family and moral; juridical and economic; mobility and political phenomena, and so on); second, that between the social and the non-social (geographic, biological, etc.) phenomena; third, the study of the general characteristics common to all classes of social phenomena."

Of these three delimited fields, the third is our chief concern for a systematic Sociology. The first field of border zones and relations between more than one, but less than all, of the social sciences, is a partial step towards the third area of studying "the general characteristics common to all classes of social phenom-

ena." This last will be our working definition of Sociology. The general characteristics of time, place, people, and the many other characteristics of people and of their environment will be the subject matter of our sociological system, regardless of whether these general characteristics appear in economic, political, religious, educational, recreational, medical, or other traditional classes of social phenomena. We do not thereby exclude the first two subfields from the general field of Sociology.

The general theory of the science of Sociology should be the system of generalizations which are general to all societal phenomena and so to all the social sciences, and should not be limited to any one or pair, or triad of them. This does not mean that Sociology is a synthesis of all the social sciences in the sense of including the content, principles, and approaches of them all. Each has its own. Sociology is viewed here as one of the social sciences and as concerned with the greatest common denominator of the others, i.e., with what is common to them all, neglecting what is specific to each. Thus, processes of co-operation, accommodation, conflict, and competition occur in economic, political, religious, recreational, or other fields of societal phenomena, and hence belong to Sociology by our definition. Similarly, such concepts as association, differentiation, density, isolation, contact, interaction, group, plurel, culture, progress, forces, and control, are common to all classes of societal phenomena.

Along with Sociology the related terms "society," "societal," and "social" need defining as to their usage in this volume. "Society" will denote human beings living together. Ordinarily as a collective noun it will connote the whole of living humanity. In the plural it denotes humanity at different periods or large relatively self-sufficient portions of it. "Group," "plurel," and "community," will be used with more exact definitions (as specified in Chapters VII and VIII) for portions of humanity, while "society" will be used as a looser, broader, but more convenient term for people-in-association without specifying how many people or how much association. "Societal" will be the adjective meaning "of, or pertaining to, associated human beings." It will be consistently used in place of "social" which will be discarded (except when quoting from current literature), because "social" denotes non-human as well as human phenomena, and because it

has become laden with many diverse and irrelevant connotations such as, "pleasant fellowship," being "extroverted," "fashionable," etc.<sup>1</sup>

In connection with defining Sociology the reader should be informed at the start that the theory developed in this volume may transcend Sociology as defined and be applicable to any or all of the social sciences. The twelve distinctive concepts of this theory can be developed for the specific content of any of the social sciences. The emphasis in this volume, however, is to develop these twelve distinctive concepts to fit the facts and concepts current in the sociological literature.

#### IV. THE CONCEPT OF "QUANTITATIVE" (IN THIS VOLUME)

A third specification of our working hypothesis after "Science" and "Sociology" was that our systematics of Sociology was to be quantitative. At first thought the distinction between the qualitative and the quantitative seems obvious, but there is a controversial border zone through which we propose to draw an exact boundary.

A continuous development of increasing precision in human observations can be described from observations that are qualitative, such as "white," "woman," "intelligently," and "co-operating," to their quantification in some way, as in degrees of whiteness, femininity, intelligence, or co-operation. Briefly, our theory of precision in measurement (which is elaborated with symbolic notation in Chapter III) is that human observation typically begins with distinguishing and naming a quality. Its presence or absence then makes that quality a variable, varying between two points, "all" and "none." Finer differentiation follows as in the use of the comparative and superlative degrees for any adjective or adverb as suggested by such terms as "more," "most," "less," "least." Qualities can be quantitatively compared as unequal before any units have been discerned; as when we judge that Miss X is less beautiful than Miss Y, or that W is pleasanter than Z. From ranking the degrees or relative amounts of qualitative entities, human observing proceeds to define standardized units wherever possible, so that ordinal series (ranks) can be converted into cardinal series, which are multiples of stand-

ardized units. Many phenomena occur with such obvious units as to be readily counted and expressed in cardinal units with little or no intellectual labor to define the standard unit. Cardinal series, in turn, are further refined in precision by calibrating them in various ways as described in Chapter III. In this series of steps from the purely qualitative to precise quantities of the qualitative,<sup>2</sup> we draw the boundary line between the qualitative and the quantitative where the constant with only one value becomes a variable with either of two values—all or none. The presence or absence of a qualitative entity is conventionally assigned numerical values of unity or zero, and this becomes a primitively observed quantity. Every percentage that has ever been computed is an arithmetic mean of such an all-or-none quantity—a fact which should assure such all-or-none variables being included in the category of the quantitative.

There is no attempt in this scheme to lump ordinal scales, such as attitude tests where the step-interval between statements of attitude may be unequal or unknown, with cardinal scales, to give an impression of exactness which has not been achieved in fact. On the contrary our S-theory explicitly provides separate notation for ordinal and for cardinal series in order to clearly classify quantified series into varying degrees of precision in quantification.

This theory of measurement,<sup>3</sup> in addition to delimiting “the quantitative,” serves to define a societal dimension—the theme of this volume as suggested in the title. A quality, as long as it remains unchanged, can be thought of as a point. It is just some one thing, representable geometrically as a dimensionless point in societal space. A point is not quantitative in itself; but let that point be taken along with another point and immediately these two points determine a line. Let the second point, for reference, be an easily ascertained one, such as the absence of that quality (which is the same as its presence in zero amount). The presence and the absence of the qualitative entity are two points determining a linear dimension. Other points may be unobservable but can be inferred to be possible.<sup>4</sup> Thus, when the frequency of occurrence of such a quality in some specified universe is observed, the percentage of actual to possible occurrences specifies an intermediate point between complete absence and

complete presence. Next, observing that quality in ordinal series, as more or less of it, specifies further points and intervals on that dimension. Finally, cardinal units specify equal and interchangeable line-sects which can be laid off end to end along that dimension and counted to measure lengths along that dimension, i.e., quantities of that quality. One can count the units, from any given point as origin, in either of two directions along the line that visualizes a dimension—hence, the name *di*-mension. A dimension may be roughly thought of as a line, or more exactly as a linear magnitude. A societal dimension represents some societal qualitative phenomenon either in its primitive form as a point (relative to a zero point which is its absence) or as a line with distances along it specified with any degree of precision. The *direction* of the line can represent (as will be elaborated in later chapters) the quality, the kind of phenomenon, while its length represents the quantity, the amount of that phenomenon. These directions and lines can be combined by addition, subtraction, multiplication, etc., by the rules of geometry applied to societal dimensions as well as to physical dimensions.

A societal dimension, then, may be roughly thought of as any societal phenomenon whose amount is determinate at least to the extent of determining its presence or absence. It is any measurable societal phenomenon. In its incipient (i.e., dimensionless) stage of observation it can be any observed societal phenomenon whatever. The fuller study of these societal dimensions, their more specific definition, their classification in an orderly system, with operational technics for determining their number, their lengths, their directions, and their combinations, in any observed set of societal phenomena, is the central theme of the ensuing chapters. (For one example, see the general formula for the number of dimensions in a situation given by Eq. 35, Ch. III.)

For an adequate exposition of the underlying postulates concerning measurement, units, and quantities in science, the reader should study again these chapters in Lundberg's *Foundations of Sociology*. (Ref. 43.)

The progress of a science is marked by the increasing precision of its data (by which we mean the recorded observations of phenomena). The prestige and the utility of the exact sciences are,

in part, a result of increasing accuracy in observing and measuring their phenomena. The steadily increasing dependence of sociologists on instruments yielding tabulatable and quantifiable data and on appropriate statistical technics for summarizing and for discovering relationships in their data is a healthy progress towards a more exact science of Sociology and away from its earlier status as a literary and philosophical discipline. As a contribution intended to accelerate this development, the S-theory presented in this volume deals with that part of societal data which is quantifiable. The theory covers much that has not yet been quantified and it has far-reaching implications for qualitative data. But its express field is societal phenomena that can be observed and recorded in quantitative form. It is thus limited to part of the field of Sociology—how large a part and how rapidly increasing a proportion remains for time to tell.

Because the field of this theory is quantifiable societal data, the quantitative may seem to some readers to be overemphasized. This may be as inevitable as that a textbook on Psychiatry seems to overemphasize the abnormal and neglect the normal which is outside its specified field. The reader should be assured that the qualitative data is not to be underemphasized. *All* data is qualitative in part. Upon the excellence of its qualitative analysis, classification, etc., the fruitfulness of all further treatment depends. Our contention is simply that to stop with qualitative data is to rest content with primitive, or incipient, observation. Qualitative phenomena need to be ever more accurately observed, yielding quantitatively observed qualities, as science progresses towards its goal of more exact prediction and control. Swift insight, inspiration, and imagination discerning new qualitative phenomena are essential at the frontier of a science, but so are plodding mathematical logic, measurement, and checking, if the ground gained is to be consolidated and permanently held. Subjective insights and objectifying technics, in order to achieve and test those insights, must supplement each other. In this teamwork, the present volume aims to contribute more to the objective and quantitative aspects of Sociology in the belief that these have been underdeveloped hitherto. This will necessarily seem to limit this volume to a part of the entire field of Sociology.

## V. THE CONCEPT OF A SYSTEM (IN THIS VOLUME)

The fourth specification of our working hypothesis calls for clarifying our concept of a *system*. A system, as suggested above in our discussion of systematizing, consists of an interdependent set of generalizations, an organization, or pattern, of principles which condition each other. In the present study our system is built out of the symbols for twelve distinctive concepts (plus the four operations of arithmetic). The combinations and permutations of these sixteen basic concepts will be shown to define by formulae more than four hundred derived concepts which summarize and comprehensively classify quantitative societal phenomena.

This systematic theory is largely a methodological one. Four of its concepts deal with content, the dimensions themselves, in naming the kinds of societal phenomena. Four other concepts (the "scripts") modify these in describing their number, precision, and classificatory relationships. The other basic concepts are "operators," stating the procedures by which the previous concepts are combined. The theory tells little of what relationships to expect between phenomena. One cannot solve for unknowns from it without further data. The equation which specifies it is a descriptive rather than a calculative one. It attempts to describe in operationally defined terms the situations *as observed*. It takes whatever data the observer records, good, bad, or indifferent, and describes in definite symbols the operational degree of precision of those data, tells how they may be classified, and prepares them in standardized and parsimonious form ready for further manipulation to discover deeper relationships in those data. The function of the theory is thus, largely, to improve methodology systematically, more than to immediately state a system of generalizations about the behavior of societal phenomena. It is a systematic way of expressing societal data, and not, directly, a system of the functionings of societal phenomena. It does not seem possible to deduce theorems and corollaries about phenomena from this theory alone. The theory simply converts societal phenomena into a *system* of data; it expresses the recorded observations of phenomena in a more orderly arrangement.

The system here lies in our organization of data, in our arrange-

ment of concepts, and is not to be thought of as "in the phenomena." The systematizing is a form of human response to stimulus situations and, as Lundberg has clearly pointed out, is not in the stimuli themselves. For this systematizing type of response to our environment the indispensable tool is the symbol. Symbols are signs enabling people mentally to manipulate phenomena whether present or absent, real or imaginary, past or future, few or many, well or ill understood, with greater power than if people were limited to responding only to phenomena which were currently stimulating their sense organs.

The chief type of symbols is words. Language, especially when written, is the most flexible and complete symbolic system man has developed; but for scientific purposes ordinary language becomes inadequate. Symbols of greater precision and objectivity are required in building an exact science. In teaching Sociology, the unstandardized and variable meanings, the emotional connotations, the profusion of terms of overlapping meaning, are bewildering to the student.<sup>5</sup> The author's experience in teaching, in English, students brought up in a dozen different languages, has emphasized the need of a more internationally standardized and precise symbolism. After ten years of experimenting with different types of symbolism, mathematical notation stands out in his experience as the most powerful, the most universal, and the most precise.<sup>6</sup>

To adapt mathematical notation to societal data the most useful branch of mathematics is statistics and an extension of matrix algebra. Since a matrix is any rectangular arrangement of numbers in rows and columns, and since almost all quantitative data in the social sciences are, or can be, so tabulated, matrix algebra is a well-adapted tool. We have extended the notation of matrix algebra with one new operational symbol (the colon) and standardized the scripts. The resulting system of sixteen symbols of an algebraic type has proved highly adequate to describe, classify, and manipulate quantitative societal data. The degree of this adequacy is evidenced by the fact that in our sample of fifteen hundred graphs, tables, maps, diagrams, formulae, and prose paragraphs, or other sets of quantitative data, like the three hundred quantitative situations reprinted in this volume culled from the postwar literature of all the social sciences, *we have not been*

able to find one single set of data that could not be described by a formula composed of those basic symbols. We are still searching for a set of quantitative societal data which cannot be reduced to a formula in terms of those symbols. One single irreducible set of data will be more valuable in challenging the adequacy of the system which we have labeled the "S-theory" (developed in Chapter II) and in leading to its improvement, than a hundred sets that can be readily reduced to an S-formula.

Somewhat offsetting, however, the scientific advantages of a more adequate symbolism for Sociology and the social sciences is the psychological disadvantage of the difficulty of learning a new language. Algebraic symbols are difficult for many people, including a large proportion of the older sociologists, to learn and to use with facility. Over and above the inertia of mastering new associations, and the effort to form new conditioned responses, there is often a residual resistance, from early schooling perhaps, to algebraic manipulations. Inferiority feelings from unsuccessful hours in mathematics classes often leads to rationalizing that mathematics is all right for the exact sciences but is inapplicable to human, or to psychic, or to social data. Resistances due to these and other causes are to be expected and the proposer of new symbols must meet it halfway, first, by making the utility or need of the algebraic symbols very convincing to the reader, and second, by making the use of the symbols very clear and simple. This we shall try to do.

To emphasize the need for better symbolism in Sociology let a few other sociologists and students of scientific methods speak. We quote from Phelps' excellent discussion of "Symbolic Sociology," in his *Principles and Laws of Sociology* (Ref. 55):\*

"It is safe to predict that sociology will be increasingly dependent upon the resources of a symbolic logic and a symbolic notation. Both of these supplements are necessary to the further study of its data as social relations or in the analysis of these data in harmony with the postulate of the interdependence of social wholes and parts. Lundberg has anticipated this need as an essential prerequisite to research in the social sciences in his statement that a standardized system of symbols is necessary if scientific facts are to be stated uniformly and intelligibly." [Ref. 45, p. 31.]

\* Reprinted by permission of Harold A. Phelps and John Wiley & Sons, Inc.

"Symbols make two important contributions to science. They are the special instruments by which man has accumulated, stored, and transmitted knowledge. Through these human characteristics, man is distinguished from animals by the designation 'the employer of signs.' The first contribution of symbolism is that human knowledge has its origins in signs, language, writing, and numerical systems. The second contribution of symbols is their aid to thought. 'Thinking,' says Eaton, 'is activity, or rather the suspension of activity, through symbols.' [Ref. 23.] Therefore, symbols perform the double function of developing and transmitting knowledge and of representing percepts (or real objects) and concepts (or imaginary objects).

"Whenever it is possible, as in mathematics, symbolic logic, and some of the physical and natural sciences, signs or formal symbols are substituted for words and the rules of grammar. In these fields, symbols have none of the subjective colorings of words. They are objective and universal in their meaning. They have precise uses and uniform interpretations. Furthermore, they are subject to verification and proof, especially when they are employed as formulae of repeating phases of reality. For these reasons, Young has concluded that 'the final arbiter of objective reality is the impersonal, non-emotional correlation of phases of reality expressed in symbols universally agreed upon.' [Ref. 83.] Valid examples of such use, however, are ordinarily limited to small areas and brief periods of time. . . . Every advance in science, according to Ogden and Richards, may be described either as the substitution of some symbolic shorthand device for a previously accepted metaphysical concept or as the clarification of mystical and confusing adhesions to symbols. . . . According to Shipman's conclusions, a symbolic sociology is both possible and necessary. It is possible whenever a science is concerned with (1) limited facts and relationships and (2) the acquisition and communication of knowledge. It is necessary because no science is better than its method of symbolic representation. Moreover, literary symbols are wholly inadequate in the task of representing the multiple and simultaneous processes with which sociology deals. [Ref. 62.]

"Consequently, a present task of sociological methodology is the organization of a system of symbolic notation. This system must be a special sociological symbolism, adaptable to the particular nature of social phenomena, and needs not be an imitation of current developments in mathematics or symbolic logic." [Ref. 55.]

The argument for the necessity of a system of more exact symbols in Sociology has been further developed by Lundberg in the preceding companion volume, especially in Chapter II. This chapter could well be quoted in full at this point in our argument.

The resistance to new symbolism is as old as the history of science. Such resistance can be defended as a useful sieve which refuses to pass unnecessary or ill-adapted symbols. To do this, however, the symbols must be studied before accepting or rejecting them. The example of Thomas Hobbes is a warning. Wallis submitted his work on Conic Sections for Hobbes to criticize. Instead of describing geometrical figures and proofs in Latin sentences as had been the invariable custom previously, Wallis introduced the innovation of lettered diagrams for which students since his time may well be grateful; but Hobbes sent it back with the comment that he knew not whether the proofs were good or bad for it was so "covered o'er with the scab of symbols" that he would not even read it! (Ref. 5, p. 427.) It is to be expected that some sociologists will react to the present volume as Hobbes did to Wallis' innovation, though opposition is more likely to be well rationalized and not so frankly based on the natural difficulty of learning something new.

History also records instances, not of apathy, but of active opposition to improved symbols. One is the invention of the zero, which was first called a "cipher" from the Arabic word and considered so cryptic ("Why name something that doesn't exist?"), that it has left its impress after almost a millennium on the English language, where, "to write in cipher" still means to write in an unintelligible language, i.e., to be "undecipherable." In medieval Marseilles the growing use of the ten Arabic digits which brought us the decimal system was forbidden by a law which prescribed penalties for the merchant who might be reluctant to do his arithmetic "in plain and clear Roman letters."

To be sure, the oblivion of many proposed symbolisms may be an excellent thing for science. But equally obviously, when Sociology needs a more exact symbolism, a proposed system should be studied carefully to determine whether it deserves oblivion or memorizing for use.

In addition to convincing the reader that better symbols are needed if Sociology is to progress towards becoming a more exact science, it is incumbent upon the proposer of new symbols to make them as simple and as clear as possible. Towards this aim the five following technics have been used in this volume:

1. The symbols have been parsimoniously adopted. Several

score have been eliminated as unnecessary, in the process of reducing their number to the sixteen basic concepts whose special cases and combinations yield all others in this volume. Of these sixteen only seven are new ( ${}^iI'_i ::$ ).

2. Familiar symbols have been preferred. Thus, of the sixteen basic symbols four are the familiar arithmetic operators of adding, subtracting, multiplying, and dividing ( $+ - \times /$ ), one is the exponent, and three are the letters T for time, L for length of geographical space, and P for the size of a population. All of our new symbols can be found on the keyboard of any typewriter. Neologisms have been almost eliminated. Even the half dozen archaic words revived here for technical use can be dispensed with by using less convenient phrases (such as "quantified indicator" for "indicant").

3. Definitions have been explicitly given for every symbol, usually in operational form of an equation in which the right-hand member states the ingredients and the procedures to calculate the entity defined. A glossary and the topical index in the appendices enable one to turn to any definition conveniently.

4. Illustrations have been liberally included. The three hundred graphs, appended to the next eleven chapters, illustrate many times over all the symbols and most of their compounds, while evidencing the extent to which this symbolic system fits the wide range of data in the literature of the social sciences.

5. Two levels of exposition have been provided, the main text and the copious footnotes. The more technical and detailed notation has been relegated to footnotes which can be skipped by the reader who wants to understand the theory but is not trying to master it sufficiently to write formulae for quantitatively recorded situations. To use the theory fully the footnotes should be studied, but to comprehend it for a preliminary appraisal they can be ignored. Almost every algebraic formula is verbalized in a sentence following it, or in the "legend" alongside each of the three hundred situations (i.e., graphs, tables, etc.), so that any reader can, with some profit, read the entire text without having to read and interpret the notation if he so chooses. It is hoped that this device will enable the non-mathematical reader to read on and understand this systematic theory, while at the same time providing the reader who grasps the notation with a correspondingly

greater power to analyze and synthesize the societal situations that are symbolized.

The specifications have now been stated which constitute our working hypothesis, that it is possible with our present knowledge to begin constructing a quantitative systematic science of Sociology. A proposal towards such a quantitative systematics is the "S-theory" outlined in the next chapter and explored sector by sector in the following nine chapters (Chs. III-XI). The last chapter attempts a review, and an appraisal of the theory as a contribution to the science of Sociology.

## VI. NOTES

1. For nine senses of the word social and for fuller discussion of these terms, "society" and "societal" see Eubank (Ref. 25, pp. 22-24 and 130-132) whose definitions we adopt here.

2. Compare Thorndike's dictum that "Whatever exists at all, exists in some amount." This basic postulate that everything qualitative is quantifiable, is more fully discussed with attendant implications in Lundberg's companion volume on postulates, *Foundations of Sociology*. (Ref. 43, Ch. II.)

3. By measurement we denote the act of determining a quantity, i.e., the ascertaining of whether a qualitative response to a phenomenon is possible as well as the ascertaining of intermediate degrees or amounts of it. The essence of measuring is *expressing in numbers*.

4. This implies non-Aristotelian logic in asserting *degrees* of such Aristotelian laws of logic, as that "A is A," "A is not B," in seeing that A may shade off on a continuum into B so gradually that any boundary is a purely arbitrary convenience and not inherent in the nature of A or B. For an exposition of non-Aristotelian logic see Ref. 36.

5. An impressive exhibit of the lack of agreement upon even the major concepts in current use in the sociological literature, together with a list of some 332 in occasional use, may be found in Ref. 25.

6. For an analysis of the levels of increasing precision (and generality at the same time) of symbols from the verbal to the mathematical equation see Chapin's "levels of symbolic substitution" (S. 1, Ch. XII) quoted and discussed by Lundberg (Ref. 43, Ch. 5).

## Chapter II

### THE S-THEORY, S

#### I. EXPOSITION OF THE THEORY

##### A. Preliminary Statement of S-theory

Towards constructing a more quantitative systematic science of Sociology, as specified in the previous chapter, a theory will be offered, which is outlined in this chapter and developed in the succeeding chapters. This theory, which will be labeled the "S-theory" from its most distinctive symbol, may now be stated. It generalizes societal phenomena in the statement:

"PEOPLE, ENVIRONMENTS AND THEIR  
CHARACTERISTICS MAY CHANGE."

This obvious generalization becomes even more obvious if the time period in which the change is observed is prolonged. Phenomena that appear static usually are seen to change if observed long enough—over millions of years for geological changes, for example. The amount of change may vary from zero up. Completely static or timeless phenomena, such as the mathematical equality  $2 + 1 = 3$ , represent the lower limit of change where its amount is zero and it passes from the category of the dynamic to the category of the static. Both categories, however, can be included in an algebraic symbol which may represent zero just as well as any positive amount of change. To achieve such increased flexibility and yet increased precision, let the principle be stated in algebraic symbols. But first, the verbal formulation of this principle may clarify it, if formulated with somewhat more care than in the seven-word sentence above.

Since our finite human minds cannot observe the totality of change of people, environments, and their characteristics, all at once, we have to observe it in portions or samples. Let any conveniently observed portion of societal phenomena be called a "situation." The record in quantitative symbols of such an ob-

served situation will be denoted by the letter S and referred to as an "S-situation." An S-situation may be thought of as any orderly set of societal data. It is any tabulation, graph, map, some diagrams, and prose paragraphs, which record societal observations quantitatively (recalling that quantitatively, as defined in Chapter I, includes frequencies of mere "presence or absence"). The three hundred S-situations appended to the chapters of this volume specify more in detail the meaning of this concept of a "situation." The boundaries defining a "unit situation" will be more sharply drawn later on.<sup>1</sup>

A preliminary formulation, then, of S-theory, to be more rigorously stated at the end of the chapter, is:

"Any quantitatively recorded societal situation can be expressed as a combination of indices of *time*, of *characteristics* of people or of their environments, of *space*, and of *population*, modified by *exponents*, and by three other scripts specifying the kind and number of *classes*, *class-intervals*, and *cases* of what the index denotes, and combined by the signs for *adding*, *subtracting*, *multiplying*, *dividing*, *aggregating*, *cross-classifying*, *correlating*, and *identifying*."

The concepts in italics in this statement constitute the sixteen basic concepts of the theory presented here. Their variations and combinations and permutations will be systematically shown in the rest of this volume. The meaning of each of these sixteen concepts will be next explained in turn.

### B. The Indices of the Four Sectors

#### 1. SPACE

The most obvious aspect of the human environment is geographic space, measurable in units of length, area, or volume. Ecological studies, recorded in part in maps of all kinds, explicitly involve this omnipresent aspect of societal environments. The man-land ratio, all travel, migrating or moving about of any kind, the effects of geographic location as a partial determinant of culture, all involve geographic space for this accurate recording. Accordingly, indices of length are adopted as one of the fundamental types of indices for describing societal situations.

## 2. TIME

All societal situations exist in time as well as in space. All action, behavior, change, dynamic phenomena, involve the temporal aspect of our total environment, and most static phenomena, explicitly or implicitly, are located and exist in time.

Of man's total culture, science especially is concerned with the time dimension, since all its laws are derived from past experience of phenomena, and all its aim of prediction and control of phenomena concerns future events. The fundamental importance of the concepts of time and space in human adjustment is such that indices of time have been adopted here as a second fundamental type of index for describing societal situations. The concepts of time and space have implications which may be perplexing metaphysically, but at the operational level of science they are readily understood and objectively determined by instruments such as clocks and meter sticks.<sup>2</sup>

## 3. POPULATION

A third fundamental type of index for describing societal situations is one which marks off the social sciences from the physical and non-human biological sciences, namely, the human population. This is the one aspect of societal situations which is "common to all classes of societal phenomena," and therefore fulfills, "par excellence," our definition of sociology (see Ch. I).

The population primarily means, in this volume, the number of human individuals, regardless of age, sex, mortality, or any other characteristic of the individual or of the plurel of individuals. Every person, or individual human being, is considered as equivalent to every other person in respect to being a carrier of characteristics. The amounts and kinds of these characteristics vary, of course, giving rise to all the observed individual differences between persons. This "pure" person, aside from his many characteristics, is never directly observed in society, but is a convenient abstraction, giving equal and objectively observable units of the population. By extension, the unit of population may be shifted from one person to one plurel, i.e., to a collection of persons. The population index in this case becomes the number of such specified plurels instead of the number of persons.

Since all human groups (which are the distinctive subject matter of Sociology) are interacting plurels, it is obvious that indices measuring their number may be a fundamental aspect of any societal situation. However, whether the unit of population is one person, or one specified plurel, the index of population can be objectively determined by the operational technics of a census.

#### 4. INDICATORS

The fourth sector, "characteristics," is defined as the societal residue. Characteristics may be anything of, or pertaining to, human beings, other than time, space, and the number of persons. They include environmental as well as physiological and psychic characteristics. They may include especially all of human culture, i.e., habitual human ways of thinking, acting, and the equipment used. This enormous field is more carefully specified in the classification of indicators offered in Chapter III. The utility of such an inclusive concept will become clearer as one studies the graphs and formulae throughout this volume. One feature of it is that it yields a perfect classification in the sense of fulfilling perfectly the canons of classification, for with this definition the four sectors are totally inclusive of quantified societal phenomena, which is the field to be classified, and are also mutually exclusive categories with no overlap and with definite boundaries.

Since the term "characteristics" includes much that can only be inferred from observation of their indicators, the term "indicator" will be used instead. Indicators are the societal facts which are objectively observable and which serve as signs of characteristics which often cannot be directly observed. Thus the behavior of a person, including speech behavior, may indicate his feelings and states of consciousness. The observable expansion of mercury in a thermometer indicates the degree of heat, a characteristic less accurately observed by the unaided senses. Again, a census may indicate the extent of unemployment characteristic of a population. Indicators may also be the unmediated characteristic itself, if the latter is in a definite and objective form such as "the number of sheep in this flock."

Usually the chief research problem is to discover or invent indicators for intangible characteristics. Where this is achieved scientific advance is made; where it remains elusive the study of that

field remains more of an art than a science. Since Lundberg has fully discussed this issue in his preceding companion volume, the scientist's faith is merely reaffirmed here that what is at present unachieved should never be considered as impossible of achievement. Negative proof is inadequate; positive proof of impossibility must be produced. To assert that certain characteristics are inherently so intangible as forever to defy reduction to objectifying indicators is the unscientific stultifying of science. To assert the possibility of developing such indicators is the only attitude which progressively leads to developing more and more of them. At a given date, however, the fact should clearly be recognized that some characteristics have objective indicators, others have none as yet, and intermediate degrees of objectivity exist. Hence, the distinction between "characteristics" and "indicators" is essential; a characteristic may be anything one can name; an indicator calls for an operational specification of the degree of its reliability and validity in representing its characteristic.

The fuller and more exact meaning of indicators, as well as of the other concepts of S-theory, are given by the three hundred graphs and their formulae in succeeding chapters. These constitute a more amplified definition of the concepts of S-theory. The sentences above are merely a verbal outline of their exact meaning.

## 5. SUBINDICATORS, E.G., "DESIDERATA"

Each of the social sciences and social disciplines specializes in studying certain societal indicators and neglecting the myriad others that exist. Thus Economics deals chiefly with the monetary indicators of exchange values. Education deals largely with the indicators of children's achievement in the activities and situations which our schools consider desirable. Every religion deals with indicators of desiderata which its devotees believe to be of supreme importance. So fundamental to many societal situations is this concept of "desideratum," the object of human desire, that it is a major subtype of indicator. Such indicators may be of positive or negative desiderata, depending upon whether people behave so as to increase or decrease the object desired. Indicators of desiderata have diverse forms and contents in the

various social disciplines, but the concept of desiderata is general to them all. Indicators of desiderata, i.e., objects of human *desire*, are the objectively observed signs of what may be denoted by such current sociological concepts as "wishes," "drives," "motivations," "wants," "appetites," "interests," "feelings," "mores," or other concepts connoting responses of approach or withdrawal. Generalized indicators of desiderata are an example of the contribution of this theory to Sociology, defined in the introductory chapter as "the study of the general characteristics common to all classes of social phenomena." The four types of indices and their subtypes, such as generalized indicators of desiderata, are claimed to be common to all classes of societal phenomena.<sup>3</sup>

#### 6. SITUATIONS—SUMMATIVE INDICES (I) AND AGGREGATIVE MATRICES, S

The indices of time and space, population and indicators of many kinds, combine in various ways and in amounts varying from zero upwards in every societal situation. In general, their ways of combining are two, namely, by summation in the broadest sense and by aggregation. Thus indices may be added or multiplied together giving a product, a sum, or a mean. Multiplication is but repetitive addition; subtraction is included in algebraic addition as the adding of negative numbers; division is a form of repetitive subtraction giving such indices as index numbers and other ratios. All four processes of arithmetic are included under summation in the widest sense (mathematicians lack a unifying label for all four).

In aggregative combinations, on the other hand, series of indices are collected and presented, usually in a tabulation or a graph. The population, or income, or other statistic, is tabled by years, regions, or other grouping. In summative combination a single variable, an index, is the result; but in aggregative combination a collectivity of indices results (such as in graphs S. 12, 29, 31). Each aggregate, represented by one of the cited graphs, may be arranged as a *matrix*, which is any rectangular arrangement of numbers in rows and columns. In societal data the matrices are sometimes imperfect, in that items for some of the cells (intersections of row and column) are non-existent or are

not recorded. Any index can be considered as the lower limit of a matrix which has but one cell in one row and one column, so that the matrix is the more general situation.

An index, or combination of indices, presented as either a single index or as a matrix, will be called an S-situation. An S-situation is any unitary<sup>4</sup> recorded societal situation, or set of recorded societal phenomena, which can be quantitatively expressed.<sup>5</sup> Each of the graphs, tables, pictures, paragraphs, or formulae in the three hundred figures in this volume is an S-situation (except where two are explicitly grouped). All these define S better perhaps than the sentence above. An S-situation may be thought of as any series of societal statistics. It includes time series and dynamic phenomena as well as static phenomena. Because of the static connotation of the term "situation," its use is not altogether satisfactory, but none of the alternative terms, such as "societal development," "societal set of data," "societal matrix," "societal statistics," or "societal pattern," seem much freer of connotations that are unsuitable for the purposes of this volume.

It is convenient to call each of the four *types* of indices composing a situation, a "sector." A situation then is analyzable into the temporal, indicatory, spatial, and populational sectors, each sector having any number of indices of that type.

## 7. SYMBOLS

The concepts of "situations" with their constituent "indices" modified by "scripts" and interconnecting "operators" will become more precise if expressed in general symbols. Such symbols avoid the irrelevant connotations, often emotional, which words have acquired and promote the objectivity of definition and mental manipulation demanded by science. They represent phenomena with greater parsimony, and parsimony is a fundamental function of science. They facilitate mathematical treatment with its resulting rigor and precision of reasoning.

Let *T* denote *time*, the length of a period, or the number of units of a duration (e.g., conception-to-birth, a transatlantic passage, 10 years, 4 working days, 5 generations, etc.).

Let *I* denote an *indicator* of some characteristic. Indicators may be qualitative or quantitative. The particular characteristic

denoted and its amount, or its units, will be specified by its scripts as explained below.

Let L denote the number of units of physical *length* (e.g., inches, kilometers, fathoms, etc.). The scripts specify the type of unit.

Let P denote the number of *persons* in a population. The characteristics identifying the persons, or the population, are expressed by indicators (or in condensed form, as explained below, by the scripts attached to the P). P will denote a population of plurels as units only when explicitly so symbolized by a script, as stated below.

Let (I) denote an *index* which combines by summation (as defined above) one or more components into a single variable (e.g., an average, a ratio, an index number, a correlation coefficient, a density, or other compounded statistic).

Let (I') denote any particular one of the indices denoted by the capital letters T, I, L, or P. This will be called the "homosectoral index," whereas without the prime the index may combine more than one homosectoral index from several sectors and thus be a "heterosectoral index."

Let S denote a *matrix*, or aggregation of indices (e.g., a table, a graph, or other unitary collectivity of statistics which can be expressed in matrix form, including the special limiting case of the matrix with one cell when the matrix becomes a single index). Thus homosectoral indices are the limiting case or simplest type of matrix. S then includes any aggregative or summative combination of indices. The S-situation is the gross unit of observation of societal phenomena developed in this volume. This gross unit of observation is analyzed into patterns of its indices with their finer units in each of the four sectors. These indices, when resynthesized according to their appropriate S-formulae, give the S-situation again.

### C. The Four Scripts

Turning from this introductory exposition of the indices, the second sentence of the preliminary statement of S-theory which deals with the scripts will be explained next.

Every index may be modified by four types of scripts, one type at each of its four corners. The scripts are ordinarily symbolized by small letters of the same kind as the capital lettered index to

which they are attached. (Thus:  ${}^tT_t$ ,  ${}^iI_i$ ,  ${}^lL_l$ ,  ${}^pP_p$ .) These four scripts are:

the exponent—(the post-superscript)	$ ^t,  ^i,  ^l,  ^p,$ or in general $ ^s$
the class script—(the post-subscript)	$ _t,  _i,  _l,  _p,$ or in general $ _s$
the class-interval script—(the pre-subscript)	${}_t , {}_i , {}_l , {}_p ,$ or in general ${}_s $
the case script—(the pre-superscript)	${}^t , {}^i , {}^l , {}^p ,$ or in general ${}^s $

A preliminary notion of each of these in turn, both in general and as applied to each of the four sectors, will be attempted next with more rigorous definition and illustration in the succeeding chapters.

### 1. THE EXPONENT, $|^s$ , DENOTING SELF-MULTIPLICATION

The exponent in mathematics denotes the number of times a factor is multiplied by itself. A negative exponent means that the factor, or base (raised to the power indicated by the exponent), is in the denominator of a ratio. Thus: <sup>6</sup>

$$\begin{aligned}
 (I)^3 &= (I) \times (I) \times (I) & (\text{Eq. 1a, Ch. II})^7 \\
 (I)^2 &= (I) \times (I) & (\text{Eq. 1b, Ch. II}) \\
 (I)^1 &= (I) & (\text{Eq. 1c, Ch. II}) \\
 (I)^0 &= (I)/(I) = 1 & (\text{Eq. 1d, Ch. II}) \\
 (I)^{-1} &= 1/(I) & (\text{Eq. 1e, Ch. II}) \\
 (I)^{-2} &= 1/(I)^2 & (\text{Eq. 1f, Ch. II})
 \end{aligned}$$

Of the fifteen hundred graphs of societal data thus far analyzed all but two or three have exponents within the range of +3 to -2.

a. *Spatial exponents:*  $L^1 = L^0$ ,  $L^{\pm 1}$ ,  $L^{\pm 2}$ ,  $L^{\pm 3}$  (Eq. 2a, Ch. II)

Situations which involve a physical or geographical length may have this represented by some index of length with an exponent of plus one,  $L^{+1}$ . If the situation expresses some other index relative to a length, as in a fare of a certain number of dollars per mile, the length is a divisor and is so denoted by an exponent of minus one. (See Eq. 1e, Ch. II.)

Recorded situations such as maps and ecological diagrams which involve areas are symbolized by an index of length with an exponent of two. In the population density, or man-land ratio, the population is divided by an area

$$PL^{-2} \quad (\text{Eq. 3, Ch. II})$$

a population density

where, as usual in mathematics, the division is denoted by a negative exponent. The sizes of this ratio and its factors distinguish urban from rural communities (as discussed more fully in Ch. VIII) and can divide Sociology into Rural and Urban subfields.

The remaining class of spatial situations involves volumes of some kind which are symbolized by an index of length cubed. Thus the gallons per capita of city water supply, or the cubic feet of room space per occupant, would be described by the ratio of a cubed length to a population, thus:

$$L^3P^{-1} \quad (\text{Eq. 4, Ch. II})$$

a volume per capita

Situations involving no physical space may be represented by a spatial exponent of zero. Strictly this represents a dimensionless space, a point in geometry, while an exponent of one represents a one-dimensional space, a line; an exponent of two represents a two-dimensional space, an area; and an exponent of three represents a three-dimensional space, a volume. But length with a zero exponent when further modified by a zero pointscript (to be explained below) symbolizes a situation without spatial dimensions, a non-spatial set of data.<sup>8\*</sup>

b. *Temporal exponents:*  $T^{\cdot} = T^1, T^0, T^{-1}, T^{-2}$  (Eq. 5, Ch. II)

The index of time whenever it states a *duration* of something, such as the age of a person or of an institution, the length of the working day, or the average duration of unemployment in a plurel, is written with an exponent of plus one. (If no exponent is written, mathematicians always understand that the exponent is +1; for examples, see S. 1 and 2, Ch. II, and all S's in Ch. IX.)

Whenever the situation involves a time *rate*, stating the velocity of any societal change, it means that the index of the amount of that change is divided by the length of time in which the change

\* For Eqs. 2b-g, Ch. II, see notes at end of the chapter.

went on. Dividing by a variable is mathematically expressible by a negative exponent. The velocity of change of some index then is symbolized by  $(I)T^{-1}$ . Consequently,  $T^{-1}$  becomes the symbol for any change in time as it is the essential and invariable factor, whether implicitly or explicitly expressed, of every situation involving change in time. To be sure, in the situation as recorded, an amount of change may be stated without stating the time period in which it occurred or the consequent time rate of change, but this is simply incomplete recording. Implicitly, if a change occurs, it has a time rate, regardless of whether the rate is variable or constant, determined or undetermined. The boundary as to when to describe the time dimension as a duration (with  $T^{+1}$ ), or as a change (with  $T^{-1}$ ), is somewhat arbitrary, and a rule for it is discussed in detail in Part V, Chapters IX and X.

For illustration, let a very simple situation be specified, as "persons forty years of age." This situation is symbolically described by the formula:

$$T^{+1} : P \quad (\text{Eq. 6, Ch. II})$$

Here  $T^{+1}$  represents the duration of life to date, and  $P$  represents the number of persons corresponding to that age. The colon is the symbol denoting this relation of correspondence (as explained more exactly below).

To illustrate change, consider the simple situation specified as "2,000,000 miles of airway flown last year." This simple but quantitatively recorded societal situation is described by the symbolic formula:

$$LT^{-1} \quad (\text{Eq. 7, Ch. II})$$

In this situation,  $L$  represents the length and  $T^{-1}$  represents its division by the time, yielding a velocity, or rate of change, in units of miles per year. Similarly, the process<sup>9</sup> of earning a living can be measured by using dollars as the indicator,  $I$ , of earnings, and a year as the unit of time,  $T$ , giving the formula:

$$IT^{-1} \quad (\text{Eq. 8, Ch. II})$$

The process of mortality is the number of persons dying per year as described in the similar formula:

$$PT^{-1} \quad (\text{Eq. 9, Ch. II})$$

The process of changing the duration of the working day, as for example, decreasing it from 10 hours' to 8 hours' duration in a decade, would be similarly described:

$$T^{+1}T^{-1} \quad (\text{Eq. 10, Ch. II})$$

This simplified illustration denotes an amount of 2 hours' change of duration ( $T^{+1}$ ) per decade ( $T^{-1}$ ). It states with algebraic generality the illustrative ratio of 2 hours/10 years, or an average rate of 12 minutes annually.<sup>10</sup>

The next class of temporal situations is that of acceleration of a societal change, described by an exponent of minus two ( $T^{-2}$ ). Whenever a change is being speeded up or slowed down the rate of change is being changed. In such situations there is a rate of change of a rate of change. This second-order change is acceleration or deceleration. It is mathematically denoted by twice dividing the index (measuring the societal change) by time. Thus the number of divorces per year is a rate of change of marital status in the population studied. An increase from one year to another in the annual divorce rate is an acceleration. The formula describing this, or any other acceleration of a population rate, is most usually:

$$T^{-1} : PT^{-1} \quad (\text{Eq. 11, Ch. II})$$

an accelerating change  
of population

The acceleration of change will be seen in Chapter XI to be the distinguishing factor in every societal force, so that the formula for an acceleration becomes the basis for measuring political, economic, educational or other societal, *forces*.

The above variation of the temporal exponent over the values of +1, -1, and -2 describes dynamic situations involving extension in time. There remain the static, or momentary, situations. These are described by an exponent of zero,  $T^0$ , as this denotes a durationless point in time. The point may be identified relative to some standard point as origin by specifying it as a date (October 1, 1938 A.D., 12:25 A.M., etc.), or it may be any indefinite point as in such a dateless assertion as, "good is the opposite of evil." Such dateless situations involve indices which

are irrespective of time and can be described by the formula:

$$T^0 : (I) \quad (\text{Eq. 12, Ch. II})$$

a dateless index, i.e.,  
irrespective of time

The subdivision of static situations into the instantaneous which exist at a particular moment and the timeless which exist at any moment of time, is specified by further scripts.

The exponent is thus seen to classify situations into the static and the dynamic, and subclassify the latter into the durative, the changing, and the accelerating.

*c. Populational exponents:*  $P^p = P^0, P^1, P^{-1}, P^2$  (Eq. 13, Ch. II)

Situations which do not involve any people can be described by a population index with a zero exponent,  $P^0$ , in combination with a zero person script, as described below under case scripts. Strictly by Eq. 1d,  $P^0$  reduces to one person and is the field of Individual Psychology rather than of Social Psychology and Sociology, which deal with more than one person. But with a zero person script the population index is asserted to be non-existent in that situation and need not even be written in the descriptive formula. Thus the time periods of the earth in S. 1, Ch. II, and the classification of the sciences in S. 5, Ch. II, are two situations involving no population:

$$(I) : {}^0P^0 = (I) \quad (\text{Eq. 14, Ch. II})$$

a populationless  
situation

This situation is a record of an index with no specified population corresponding to it—a populationless situation.

When a population, either of persons or of plurels, is specified in the situation as recorded, the homosectoral index,  $P^{+1}$ , written in the formula represents that population. When the situation is expressed in per capita units or other similar units relative to a population, it means that some index has been divided by a population. This ratio is expressed by a negative exponent:

$$(I)P^{-1} \quad (\text{Eq. 15, Ch. II})$$

a per capita index

The exponent of 1, either positive or negative, identifies the class of situations involving a plurel, a human population.

The third class of situations identified by the populational exponent is that of groups in the sociological sense of a plurel of *interrelated* persons. To show the structure of a group systematically, arrange its members in a matrix as column headings and again as row headings, and in each cell enter the index expressing the relation between that pair of persons. The aggregate of all such relations constitutes the group, as demonstrated more rigorously in Chapter VII. This matrix has  $P^2$  relations of one kind in the cells which are  $P \times P$  in number. Hence the population index with an exponent of plus two becomes the symbol (along with other qualifying scripts) of a sociological group:

$$P^2 \quad (\text{Eq. 16, Ch. II})$$

a sociological group

Thus, according to whether the populational exponent is 0, 1, or 2, the situation is classified as involving no population, plurels, or groups.<sup>11</sup>

$$d. \text{ Indicatory exponents: } I^1 = I^0, I^{\pm 1}, I^{+2} \quad (\text{Eq. 17, Ch. II})$$

The large field of indicators of the myriad societal characteristics can be neatly classified by the indicatory exponents on the basis of the thoroughness with which those characteristics have been observed and recorded in the indicator. Characteristics may be observed qualitatively as attributes, or named kinds of things; or they may be more thoroughly observed quantitatively as amounts or degrees of such attributes; or finally, they may be still more thoroughly observed relationally, as correlations of quantified attributes, measuring their interworking patterns. An attribute can be symbolized by an indicator with an exponent of zero (and a class script to be described below, specifying the quality). An exponent of one represents a quantitative indicator. An exponent of plus two represents correlated indicators.

A zero exponent means unity (see Eq. 1d, Ch. II) of whatever kind of thing is denoted by the class script. A zero exponent reduces the characteristic to its pure quality undifferentiated into quantities of it.

$I^0$  = an attribute, a unit quality (Eq. 18, Ch. II)

$I^{+1}$  = a variable quantity of some quality, an amount of an attribute (Eq. 19, Ch. II)

For an example, consider a qualitative sensation such as the color blue. This attribute can be more thoroughly observed and classified into degrees by such diverse units as those of physical frequency of light waves, chemical proportions of standardized pigments, or psychological just-perceptible-differences of either hue or saturation. Similarly, a qualitative attitude such as "hate" can be more exactly observed and classified into degrees of hatred by means of a verbal attitude test, or some other standardized scale of behavior items. Further examples of varying complexity and varying thoroughness of observation of societal characteristics are given in graphs S. 5 and 6, Ch. II, and the S's of Chapter III.

This principle, that the qualitative and the quantitative can both be represented by one symbol, I, and distinguished by exponents of zero or greater, will be referred to as our *attribute hypothesis*. It is one of the most far-reaching and fruitful hypotheses in the system of hypotheses that is the S-theory. As will be seen in more detail in Chapter III, it adds a large section of qualitative phenomena to the quantitative phenomena for which mathematical technics have proved such a powerful tool in science.

Thus by logical extension of the usual rules of mathematics, addition, subtraction, multiplication, and division of qualities by each other and by quantities, become possible. For illustration, addition of qualitative phenomena consists in following a classification table upwards, so that a class is the sum of its subclasses. In S. 5, Ch. II, Science is the sum of the five classes of science, each of which is in turn a sum of its subclasses. For qualitative division, see S. 14, Ch. II, where a pictorial analogies type of intelligence test illustrates an equation of ratios of qualities. Multiplying a quality times a quantity is modifying or "qualifying" the quantity. Thus a population, P, if multiplied by the attribute negro ( $I^0$ ), gives the product  $I^0P$  = a negro population. This extension of the logic underlying mathematics to qualitative indicators is further developed in Chapter III

and throughout this volume, but is briefly noted here to suggest the far-reaching importance of the use of the indicatory exponent of zero.

The third stage of thoroughness of observing societal characteristics is represented by an indicatory exponent of plus two which symbolizes a correlation in some form. The typical operational definition of correlation is the Pearson product-moment formula for computing the correlation as a precise coefficient. This formula involves multiplying two indicators together (when each is converted into standard deviation units) and averaging the products. It is thus a mean of the products of two quantified characteristics. In vectorial algebra this is called the scalar product and denoted by a heavy dot. This symbol is used in S-theory to denote correlation, which is a particular type of multiplication. (Later it will be shown that this scalar product is more general than the ordinary arithmetic product, which is but the limiting case of scalar products.) Therefore, the formula for correlation is:

$$I_{i'}^{+1} \bullet I_{j'}^{+1} = r_{ij} \quad (\text{Eq. 20a, Ch. II})$$

a correlation of indicators

where the primes in the subscript denote two different indicators, such as, for example, the intelligence test scores and the monetary incomes of the persons in a plurel. By the usual algebraic rule that, to multiply, the exponents are added (see Eq. 1, Ch. II), in Eq. 20 the indicator is raised to the second power (neglecting scripts other than the exponent) and is symbolized by:

$$I^2 \quad (\text{Eq. 20b, Ch. II})$$

a correlation (considering  
exponent scripts only)

Correlation is one of the most powerful tools the social sciences possess for measuring relationships of phenomena and for enabling prediction with known degrees of probability. To the degree that the correlations of a phenomenon with all other phenomena become known, that phenomenon becomes proportionately well understood for the purposes of science. Phenomena are most thoroughly observed when their correlations, the patterns of their interworking with other phenomena, are observed and recorded.

Thus it is seen that the three indicatory exponents of 0, 1, and 2 serve to classify the multitude of societal characteristics into the qualitative, the quantitative, and the correlated, which are seen as three degrees of increasing thoroughness of man's responses to his environment.<sup>12,13</sup>

*e. The quantic formula and the quantic number, |<sup>s</sup>*

In discussing indices with various exponents it is convenient to use the terms:

"Nullary" to denote any index with a zero exponent,  $(I)^0$

"Primary" to denote any index with one as exponent,  $(I)^{+1}$

"Secondary" to denote any index with two as exponent,  $(I)^{\pm 2}$

"Tertiary" to denote any index with three as exponent,  $(I)^{\pm 3}$

The first step in analyzing any societal situation, S, is to determine the exponents in each of the four sectors. The formula for S which specifies only the indices and their exponents and neglects the other three scripts is called the *quantic formula*, as quantic is the term mathematicians use to denote the degree of an equation.<sup>14</sup> When the quantic formula is written as a four digit number, specifying the exponents in order  $|^{t;i;l;p}$ , it will be called the *quantic number*. This is denoted by the symbol  $|^s$ , which means the aggregation of the exponents of matrix S.<sup>15</sup>

$S^s = T^t;I^i;L^l;P^p$  = the quantic formula of S-theory

(Eq. 21a, Ch. II)

$|^s = |^{t;i;l;p}$  = the quantic number classifying societal situations

(Eq. 21b, Ch. II)

The quantic number is the heart of the S-theory. It serves to classify every S that can exist into a single, definite, unambiguous category according to its combination of indices and exponents. *The quantic number thus provides a thoroughgoing basis of classification for all quantifiable societal phenomena.* This assertion that every quantified societal situation can be represented by a quantic number is our "quantic hypothesis."

The utility of any classification is relative to some purpose and no classification is perhaps useful for all purposes. The claim made here is, that for the purposes of a *systematic science of Sociology*, the quantic classification is a fundamental one. It



$P^{\pm 1}$	$I^0$	$P_P$	PLURELS Ch. IV	e;0;e;1	0;0;0;1	0;0;2;1	1;0;0;1	9;0;0;1	8;0;0;1	Terna- ries
$P^{\pm 1}$	$I^{\pm 1}$	${}_i(I)^{+1} : P^{+1}$	DISTRIBUTIONS Ch. V	e;1;e;1	0;1;0;1	0;1;2;1	1;1;0;1	9;1;0;1	8;1;0;1	
	$I^{+2}$	$(I), :: (I)_{//} : P^{\pm 1}$	CORRELATIONS Ch. VI	e;2;e;1	0;2;0;1	0;2;2;1	1;2;0;1	9;2;0;1	8;2;0;1	
	$I^0$	${}^pP :: {}^pP : (I)^0$	INTERRELATIONS Ch. VII	e;0;e;2	0;0;0;2	0;0;2;2	1;0;0;2	9;0;0;2	8;0;0;2	
$P^{+2}$	$I^{\pm 1}$	${}^pP :: {}^pP : (I)$		e;1;e;2	0;1;0;2	0;1;2;2	1;1;0;2	9;1;0;2	8;1;0;2	
	$I^{+2}$	$(I), (I)_{//} : {}^pP :: {}^pP$		e;2;e;2	0;2;0;2	0;2;2;2	1;2;0;2	9;2;0;2	8;2;0;2	

\* Several types of situations are omitted from the table for simplicity. Among such are:

(a) Rare ones:  $I^3$  (as in skewness),  $I^4$  (as in kurtosis),  $T^{-3}$  (as in power),  $L^{-1}$  and  $L^{-3}$  (linear and cubical phenomena), etc.;

(b) Variant subtypes: thus variance,  $\sigma^2$  is the special case under correlation where  $(I)_{//} \equiv (I)_{//}$ ;

(c) Repetitive types: thus Densities, the situations involving physical space,  $L^1$ , are omitted as a subcolumn under the  $T^1$  columns to the right of the  $T^0$  column. These  $L^1$  subcolumns exist, although not shown in this table.

† The four digits of the "quantic number" represent the exponents of the four indices, T, I, L, P, of the quantic formula for all quantitative societal phenomena. Note that e means any value, 9 means -1, 8 means -2.

may come to play a role for the social sciences comparable to the classification of the chemical atoms in Mendelyev's periodic table. Evidence as to the validity and utility of this quantic hypothesis will occupy most of the space of subsequent chapters. The titles and sequence of these chapters follow the categories of this quantic classification. The quantic classification can be set forth in a table.<sup>16</sup> (See pp. 42 and 43.)

*f. The quantic table, classifying societal phenomena*

In this table expanding the quantic formula, the main columns classify the temporal indices as nullary,  $T^0$ , primary,  $T^{\pm 1}$ , and secondary,  $T^{\pm 2}$ , or into static situations, change, and acceleration, i.e., forces, respectively. The three main rows classify the populational indices into a nullary row,  $P^0$ , where population is absent; a primary row,  $P^{\pm 1}$ , dealing with characteristics and changes of a population; and a secondary row,  $P^{\pm 2}$ , dealing with the interrelations to each other of all the persons in groups. The indicatory indices, instead of being represented properly in a third dimension by nullary, primary, and secondary pages, are arranged as three subrows giving qualitative,  $I^0$ , quantitative,  $I^{\pm 1}$ , and correlated,  $I^{\pm 2}$ , characteristics, respectively. Finally the spatial indices, which are properly a fourth variable needing a fourth dimension, are represented by subdividing the time columns. For simplicity, however, only the static column of nullary time is subdivided, and it is subdivided only for area,  $L^{\pm 2}$ , as simple length and volume,  $L^{\pm 1}$  and  $L^{\pm 3}$ , are of much less frequent occurrence in societal situations.

*The cell entries state the quantic numbers which classify any and every quantified societal situation into some one cell (or to a group of adjacent cells if it is a composite situation).*

(The columns headed "Typical Descriptive Formulae" may be ignored at present until explained further on in the section on the other scripts.)

The whole table shows an orderly progression from greatest simplicity with a nullary quantic in the upper left corner ( $|^s = 0;0;0;0$ ) to greatest complexity with a sixth degree quantic in the lower right corner ( $|^s = 8;2;0;2$ ) which, if the length component were fully represented, would be a ninth degree quantic ( $|^s = 8;2;3;2$ ). As will be more fully explored in Chapter III,

this roughly is the progression in human thinking which started with primitive observation of *qualities*, moments of time, and individuals ( $e = 0$ ); progressed to observing *quantities* of these with ever increasing precision and range ( $e = 1$ ); and finally, with modern science, is progressively reducing their *relationships* to measurement ( $e = 2$ ). The diagonal zones running from lower left to upper right show a progression, which is perfect in the true quantic solid, both in absolute size of the exponents and in number of sectors from none to all four. These zones, marked by the dashed lines and named in the last column, are imperfectly represented on a plane such as the table on pages 42 and 43.

The forty-five cells shown in this table are misleading in that the space components are imperfectly represented. The number of cells, i.e., quantic classificatory categories, that are of common occurrence, is eighty-one. This allows for exponents of 0, 1, or 2 on each of the four components ( $3^4 = 81$ ). If the rarer situations with exponents of three were included, the number of possible categories would be two hundred and fifty-six ( $4^4 = 256$ ), but for most of these categories no phenomena belonging in them have as yet been observed. They are readily conceivable, and with the progress of quantitative research in the social sciences they may be expected to be reported in the literature in the future. This table specifying their properties may even serve as a stimulus to their discovery. For example, the quantic number 8;2;0;2 identifies a change in the correlation of two dynamic traits which exist only as interrelations between pairs of persons, such as husband and wife. Such a situation has never been reported in the social sciences as yet for lack of adequate instruments of observation, but its properties are predicted by this quantic number.

The topics in the fourth column of the table are the chief concepts naming the commoner combinations of sectoral indices and their exponents. Those in capitals will be the headings of the ensuing chapters of this volume. The chapter sequence starts from the most elementary situation ( $|^s = 0;0;0;0$ ) and cumulates indices down the static column in developing Indicators  $I^{0;\pm 1}$ , then, adding population in Part III, Plurels  $P^{\pm 1}$  are developed, then Distributions  $I^{\pm 1}P^{\pm 1}$  (with social problem plurels as a minor subclass), then Correlations  $I^2P^{\pm 1}$ , analyzing the pattern of char-

acteristics, and next Interrelations  $I^{+1}P^{+2}$ , analyzing the patterns of human groups. Then taking up the columns from left to right, the space index is explored in Part IV in static situations,  $T^0$ ;  $I^i$ ;  $L^{1 \neq 0}$ ;  $P^p$ , with its definition of density delimiting rural and urban Sociology. Next, Part V takes up the time index beginning with Change,  $T^{\pm 1}$ , in its slower Duration form,  $T^{+1}$ , and in the form of Processes,  $T^{-1}$ . Accelerations,  $T^{-2}$ , are then taken up with their chief subtypes of societal Forces,  $T^{-2}IP$ , and societal Control,  $T^{-2}IP^2$ .

Further exposition of the exponent script will be deferred to later chapters. The above notion of its importance in providing a complete and rigorous classification for quantified data in the social sciences is perhaps adequate for the preliminary study of this chapter opening up the whole of S-theory, before exploring its ramifications section by section. The other three scripts will be explained next.

## 2. THE THREE AGGREGATIVE SCRIPTS = "DESCRIPTS" = $\begin{smallmatrix} s \\ s|s \end{smallmatrix}$ (Eq. 22, Ch. II)

The three scripts other than the exponent are aggregative, that is, they denote the number of entities that are collected together and can be arranged in a matrix, such as ordinary statistical tables. Because of this common property, a term to distinguish them from the multiplicative exponent script is convenient and the archaic term "descriptor" is here revived for technical purposes, connoting "that which describes." A descriptor symbolizes the number and kind of a set of aggregated, i.e., separately listed, entities.

$$a. \text{ Class scripts: } (|_t;|_i;|_1;|_p) = |_s \quad (\text{Eq. 23, Ch. II})$$

The post-subscript, which will be called "class script," specifies the number of aggregated indices of a specified kind, or class, in the situation analyzed. Thus, if the situation is a college population tabulated into classes as:

$$\left\| \begin{array}{l} 500 \text{ Freshmen} \\ 400 \text{ Sophomores} \\ 350 \text{ Juniors} \\ 300 \text{ Seniors} \end{array} \right\| = P_p \quad (\text{Eq. 24a, Ch. II})$$

illustrating plurals,  
i.e., populational  
classes

this simple matrix (denoted by enclosure between double lines)

of four rows and one column is symbolized in S-notation by a population index,  $P$ , denoting the number of persons in each of the classes,  $|_p$ , which in this illustration are four classes of college students. If, next, this population were recorded as classified by sex in a matrix of one row and two columns:

$$\begin{array}{cc} \text{Men} & \text{Women} \\ || 860 & 690 || = P_a \end{array} \quad (\text{Eq. 24b, Ch. II})$$

a second letter (by convention, the next letter in the alphabet), as class script, would denote these two sex classes. Here  $|_a$  denotes two sex classes and  $P$  is a variable with two values, namely 860 and 690 persons. Since the classes of a population are called plurels in Sociology, the populational class script can be thought of as a "plurel script," representing the number and nature of the plurels.

To give an example of a two-dimensional space in the space sector, the class script symbolizes the number and nature of areas. Thus the matrix of ecological data of a city's zoning would be symbolized as follows:

$$\begin{array}{|l|l|l|l|} \hline \text{Residential zone} & 17 & \text{square miles} & \\ \text{Business zone} & 6 & " & " \\ \text{Industrial zone} & 2 & " & " \\ \hline \end{array} || = L_1^2 \quad (\text{Eq. 25, Ch. II})$$

illustrating regions,  
i.e., areal classes

where  $L^2$  denotes the area (here in units of square miles) in each of the  $|_1$  regions, i.e., classes of area (here three urban zones).

In the indicatory sector the class script symbolizes the number and nature of the indicators in the situation as recorded. Thus in a study of national trends at some date the indicators collected in the record might be:

$$\begin{array}{|l|l|} \hline \text{Cost of living index number 105} \\ \text{School budget index number 108} \\ \text{Pages of print index number 112} \\ \text{Party in power} & \text{Democratic} \\ \text{Labor legislation} & \text{Wages and Hours} \\ & \text{Bill passed} \\ \text{Weather} & \text{Drought season} \\ \hline \end{array} || = I_1^{1.0} \quad (\text{Eq. 26, Ch. II})$$

illustrating indicators,  
i.e., classes of  
characteristics

Here the indicatory class script (or "indicator script" as it may be conveniently called) represents six different indicators of societal characteristics of which the first three are quantitatively

recorded, and are so denoted by the exponent of 1, and the last three are qualitatively recorded, and are therefore so denoted by the exponent of 0. This matrix of order  $6 \times 1$  (6 rows and 1 column) could be sectioned, or split, into two matrices, each of order  $3 \times 1$  with their respective formulae of  $I_i^{+1}$  and  $I_i^0$ , where  $|_i = 3$  and  $|_i = 3$ .

In the temporal sector, the class script symbolizes the number and nature of the different *sets* of units used for measuring time in the recorded situation. Thus the length of the working day, measured in hour units, could be symbolized by  $T_t$ , while the series of years for which such data are recorded could be symbolized by  $T_u$  (and more precisely specified by the further scripts of primes to be expounded below).

In sum, for all the sectors the class script tells the number and the kind of the different *dimensions* in the societal situation as recorded. According to whether the letters are t, i, l, or p (or letters immediately following these in the alphabet), the dimensions are of the temporal, indicatory, spatial, or populational types, i.e., sectors. The symbol  $|_s$  denotes the class scripts in all four sectors; it represents the aggregation, or pattern, of the four homosectoral class scripts, as specified in Eq. 23, Ch. II. Recall that a dimension, visualizable as a geometric line, means any quantity of something and includes the primitive two-point dimension of the presence-or-absence of that something.

The capital letter primary index to which the class script is attached states the length along each dimension in some sort of units. The exponent asserts the number of times a dimension of one kind is repeated in multiplying it by itself. (The properties of the exponential and class script types of dimensions will be distinguished below when describing the geometrical interpretation of S-theory.)

For some examples of class scripts in more complicated situations note the following:

In S. 6, Ch. II there are six indicators of "home status," i.e., six classes of the indicatory type of index ( $|_i = 6$ ). In S. 2, Ch. II, there are two time indices, as time is classified in two ways: as hours of the working day, and as years of a century ( $|_t = 2$ ). In S. 11, Ch. II, one population is classified in two ways: once into 38 national classes or plurels; and again into

21 religious classes. Here there are two indices of the population subdivided into 59 indices ( $|_{\Sigma p} + |_{\Sigma q} = 59$ ). In S. 9 the map of New York State shows two classes of area and two classes of lines, i.e., there are two secondary indices ( $L_{i=2}^2$ ) and two primary indices ( $L_{i=2}^{+1}$ ). Of the areal classes one is subclassified into counties and another into 10-mile strip and non-strip subclasses. The two linear classes are those of boundaries and of roads, which are subdivided into three subclasses.

These illustrations show clearly that a situation may have more than one index in any one sector. Since scripts are ordinarily homosectoral, i.e., attached to indices of the same letter, either the letter of the script or of the index serves to specify the sector. The symbol  $|_s$  denotes the aggregation, or pattern, of class scripts in the four sectors.

*b. Class-interval scripts:*  $(\iota|;i|;i|;p|) = s|$  (Eq. 27, Ch. II) <sup>17</sup> \*

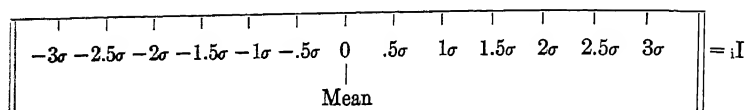
The pre-subscript is used in S-theory to specify the number and nature of the aggregated class-intervals, or the series of equal units, in which the index is expressed. Thus the series of periods (which are class-intervals of time)

$$\left\| \begin{array}{l} 0-19 \text{ years} \\ 20-39 \text{ " } \\ 40-59 \text{ " } \\ 60-79 \text{ " } \\ 80-99 \text{ " } \end{array} \right\| = \iota T \quad \begin{array}{l} \text{(Eq. 29, Ch. II)} \\ \text{illustrating periods,} \\ \text{i.e., temporal} \\ \text{class-intervals} \end{array}$$

are represented in the formula by  $\iota|$ , representing five periods, and T, representing twenty years, so that their combination is an aggregation, or series, of five twenty-year periods. This is an essential part of the formula for any time series, such as tables or graphs of societal change, population pyramids, etc.

The class-intervals of quantitative indicators <sup>18</sup> are illustrated by any frequency distribution. For example, consider a distribution of socio-economic status as measured by the Chapin-Leahy scale. If the points of score, I, were grouped into class-intervals of one half a standard deviation ( $.5\sigma$ ) each, and if the range were six sigma, there would be twelve class-intervals.

\* For Eq. 28, Ch. II, see notes at end of the chapter.



(Eq. 30, Ch. II)  
illustrating a scale  
of indicatory  
class-intervals

In this graphed matrix of order  $1 \times 12$  (1 row by 12 columns) the symbol  $iI$ , represents the 12 class-intervals, and the indicator  $I$  represents the number of units,  $.5\sigma$  in each, so their combination  $iI$  represents the complete scale which is an aggregation, or series, of class-intervals of specified size and number. Wherever no grouping of units is made the unit is identical to the class-interval. Thus the class-interval script may be used to specify the unit.<sup>19</sup>

For the spatial indices the class-interval script similarly denotes the number of aggregated quantitatively equal and interchangeable class-intervals or units which may be line sects if the index is  $L^1$ , or areas, if the exponent is  $L^2$ , or volumes, if the exponent is  $L^3$ . Note that the class-interval script, like the other descripts, denotes an aggregation not a sum; it represents the series of separate class-intervals, not a single number expressing their total extension; it requires a tabulation, a frequency distribution, a graph, or a map to show the several class-intervals collected together, whose aggregate is symbolized by the descript. Thus the formula for a foot rule of twelve inches is:

$$iL^{+1} \quad (\text{Eq. 31a, Ch. II})$$

illustrating a linear  
scale, i.e., linear  
class-intervals

which is general to any line, straight, curved, or irregular, subdivided into equal sects. The similar formula for an area graphed by co-ordinates in equal units is:

$$iL^{+2} \quad (\text{Eq. 31b, Ch. II})$$

illustrating an areal  
scale, i.e., areal  
class-intervals

while for volumes such as carloads of freight, bottles of milk, cubic feet of storage space, it is:

$$iL^{+3} \quad (\text{Eq. 31c, Ch. II})$$

illustrating cubic  
class-intervals

In the populational sector, the class-interval script conveniently denotes the number and nature of the plurels of equal size which are the class-intervals, or the units, in which that population is measured. Thus a survey may take the family as the population unit, and consider all families as of the same size, i.e., as representable by the average sized family. The formula for any such survey using families, athletic clubs, church congregations, or any other plurels, as the unit, is:

$$|| \text{ a list of interchangeable plurels } || = {}_pP \quad (\text{Eq. 32, Ch. II})$$

illustrating plurel units,  
i.e., populational  
class-intervals

In sum, the class-interval script specifies an aggregation, a series of qualitatively identical and quantitatively equal subdivisions of the index to which it is attached. The class script in contrast specifies an aggregation of qualitatively different subdivisions of its index. The class descript describes qualitative subclasses; the class-interval descript describes quantitative subclasses of one quality. The class script tells the number of dimensions, and the class-interval script tells the number of subdivisions along one dimension.

This unifying of periods of time, class-intervals of scales measuring miscellaneous societal characteristics, line sects, unit-areas, unit-volumes, and unit-plurels, all under one higher order concept and symbol,  ${}_s|'$  (the homosectoral class-interval script), is a useful device furthering the scientific aim of parsimony, which is to standardize fewer responses to an ever increasing number of diverse stimuli, and thereby, to increase finite man's adequacy of adjustment to his multiform environment. All the sixteen basic symbols of S-theory which are here being expounded, perhaps in such detail as to bore some readers, have this eventual aim of parsimony as their justification.

$$c. \text{ Case scripts: } ({}^i|;{}^i|;{}^i|;{}^p|) = {}^s| \quad (\text{Eq. 33, Ch. II})$$

The pre-superscript in S-theory specifies the number and nature of aggregated specific cases, or points, on the index which it modifies. This is best clarified by again discussing each sector in turn.

In the time sector, the case script specifies the dates, which are points in time. It is therefore referred to hereafter as the

"date script." Thus a series of financial statements, or any other data, observed on a series of dates, such as:

$$\left\| \begin{array}{l} \text{July 1, 1930} \\ \text{July 1, 1931} \\ \text{July 1, 1932} \end{array} \right\| = {}^vT \quad (\text{Eq. 34, Ch. II})$$

illustrating dates,  
i.e., points in time

would be symbolically represented in part by the date script,  ${}^v|$  (here  ${}^v| = 3$ ), on the temporal index,  $T$ , which specifies the time interval between each pair of successive dates. ( $T$  in  ${}^vT$  is thus a variable with  $t-1$  values.)

In the indicatory sector, the case script specifies particular points on the scale measuring some societal characteristic. Thus a mean, a standard deviation, the limits of the range, are points commonly specified by this script. A set of tests with their lower and upper limits can be symbolized by:

<i>Lower</i>	<i>Upper</i>	
$\left\  \begin{array}{ll} \text{Test A} & 48 \end{array} \right\ $	$\left\  \begin{array}{ll} 96 \end{array} \right\ $	$= {}^{a:z}I_i = {}^iI_i \quad (\text{Eq. 35, Ch. II})$
$\left\  \begin{array}{ll} \text{Test B} & 0 \end{array} \right\ $	$\left\  \begin{array}{ll} 50 \end{array} \right\ $	
$\left\  \begin{array}{ll} \text{Test C} & 120 \end{array} \right\ $	$\left\  \begin{array}{ll} 178 \end{array} \right\ $	

illustrating limits,  
i.e., extreme in-  
dicatory points

Here as usual, in place of the script of the same letter as its index, more specific letters may be substituted. The letters  $a$  and  $z$  are frequently used in  $S$ -theory to denote the first and the last limits within which data are presented. This is an essential and much ignored practice in the social sciences. Many contradictory and confusing correlations and other findings are reported without specifying the limits of the variables, or recognizing that what holds true in one range may not hold true between other limits. The case script, in explicitly calling attention to the limits, should help to reduce such confusion.

In the spatial sector, the point script specifies

- the number of points unlocated in space  $= {}^1L^0$  (Eq. 36a, Ch. II)
- the number of points located on a line  $= {}^1L^{+1}$  (Eq. 36b, Ch. II)
- the number of points located in an area  $= {}^1L^{+2}$  (Eq. 36c, Ch. II)
- the number of points located in a volume  $= {}^1L^{+3}$  (Eq. 36d, Ch. II)

Of these, the points on a map,  ${}^1L^{+2}$ , such as cities, stations, or identified spots of any kind located by two co-ordinates, are frequent examples.

In the population sector, the case script represents a list of

identified persons, as a person represents a point on a populational dimension:

$$\left\| \begin{array}{l} \text{Mr. Smith} \\ \text{Mrs. Jones} \\ \text{Mr. Black} \\ \text{Miss Doe} \end{array} \right\| = {}^p P \quad \begin{array}{l} \text{(Eq. 37, Ch. II)} \\ \text{illustrating persons,} \\ \text{i.e., populational} \\ \text{cases} \end{array}$$

In sum, the case script, or point script, specifies an aggregation of points on the index it modifies.

For further examples of all the descripts, almost any of the situations appended to this and the other chapters may be studied.

*d. The sectoral S-formula:  ${}_s(I')_s^s$*

The detailed exposition of the four scripts in each of the four sectors may be summarized and more easily remembered in the following "sectoral S-formula" which states the symbolic form and the meaning of each script in each sector, and in general:

$${}_s(I')_s^s = \left\{ \begin{array}{ll} {}^{20} \text{cases} = {}^s & {}^s = \text{exponent} \\ & (I') = \text{amount of the index} \\ \text{class-intervals} = {}^s & {}^s = \text{classes (qualities)} \\ \\ \text{dates} = {}^t & {}^t = \text{temporal exponent} \\ & T = \text{duration} \\ \text{periods} = {}^t & {}^t = \text{kinds of time units} \\ & \text{or} \\ \text{points} = {}^i & {}^i = \text{indicatory exponent} \\ & I = \text{amounts of an indicator} \\ \text{class-intervals} = {}^i & {}^i = \text{classes, or indicators} \\ & \text{or} \\ \text{points} = {}^1 & {}^1 = \text{spatial exponent} \\ & L = \text{length} \\ \text{quantitative} = {}^1 & {}^1 = \text{qualitative classes of} \\ \text{subspaces} & \text{space} \\ & \text{or} \\ \text{persons} = {}^p & {}^p = \text{populational exponent} \\ & P = \text{number of persons} \\ \text{quantitative} = {}^p & {}^p = \text{qualitative plurals} \\ \text{subplurals} \end{array} \right.$$

(Eq. 38, Ch. II)

Before concluding the explanation of scripts and proceeding to the operational symbols, it should be noted that the notation of scripts developed here can be entirely consistent with the usual notation of statistics. By keeping capital letter scripts to specify indices and small letter scripts to specify matrices, all the usual formulae of statistics can be written in S-notation. S-notation is limited (as here defined) to societal phenomena, whereas statistical formulae are general to astronomy and other fields as well. Statistical formulae are more exclusively concerned with mathematical forms, while S-theory combines these with sociological content, using these forms for building highly systematized content.

#### *D. The Operators*

The preliminary statement of S-theory at the beginning of this chapter involved *indices* and *scripts*, combined by *operators*. Besides the usual mathematical operators,  $+$ ,  $-$ ,  $\times$ ,  $/$ , which add, subtract, multiply, or divide the symbols before and after them, four operators, much used in S-theory, are the symbols denoting: (1) aggregating or subclassifying entities, (2) cross-classifying them, (3) correlating them, and (4) identifying single entities. These symbols are: the colon ( $:$ ), the double colon ( $::$ ), the heavy dot ( $\bullet$ ), and the primes (' ' ' ' ' '), respectively. Between them they extend to the peculiar needs of the social sciences the precision, logical consistency, and objectivity of the conventional mathematical operators with their operational rules.

##### 1. THE AGGREGATING OR SUBCLASSIFYING SIGN ( $:$ )

Entities that are collected together into a set, but with their entities preserved, are called by mathematicians an aggregation. We adopt the colon as the sign of aggregating. Every statistical tabulation, every graph, is an aggregation. But aggregations are usually arranged in some kind of order such as a matrix (rows and columns); a hierarchy of classes and subclasses, as in a diagram of a family tree, an organization chart, or a tennis elimination tournament; or, a spatial arrangement, as in a map. This implies that the entities in an aggregation may themselves be broken down into subaggregations, or conversely, that an aggre-

gation may become one entity in a larger aggregation. Thus a league of nations' organization is an aggregation of nations, which are aggregations of states, provinces, colonies, or other subplurels, as generalized in the formula:

$$P_p':q:r \quad (\text{Eq. 39, Ch. II})^{21*}$$

illustrating a hierarchy  
of plurels, i.e., a popula-  
tional classification

Here the population,  $P$ , is asserted to be of one particular plurel,  $|_p'$  (denoted by the prime), which is subclassified (as denoted by the colon) into an aggregation of  $|_q$  plurels, *each* of which is further subclassified (as denoted by the second colon) into subaggregations of  $|_r$  plurels (where  $r$  has different values,  $q$  in number).

This sign of subclassifying is used with all three descripts as well as between indices. Thus in the class-interval script, a calendar dividing a year into four seasonal periods and subdividing these time units into twelve months would be an example, symbolized in general by algebraic letters, as:

$$t':u:wT \quad (\text{Eq. 41, Ch. II})$$

illustrating hierarchy  
of regular time  
periods

This formula is verbalized as, "There is one time period,  $t'T$ , which is subclassified ( $:$ ) into a series of subperiods,  $|_u$  in number, to each of which there corresponds ( $:$ ) a series of subperiods,  $|_w$  in number." (See S. 1, Ch. IX.) Note that the colon can be put into words by such phrases as, "... each subclassified into ...," "each with corresponding ...," "... each with a subaggregation of ..."

A frequency distribution serves as an example of the use of the colon sign between indices. Suppose the socio-economic statuses of a hundred persons were distributed as shown on page 56. As in Eq. 30, Ch. II, the symbol  $|_I$  represents the twelve class-intervals of the indicator. The colon asserts that for each class-interval there corresponds a population frequency,  $P$ .  $P$  is thus a variable having values,  $|_i$  in number. Frequency distributions are one of the commonest forms of recording quantitative data

\* For Eq. 40, Ch. II, see notes at end of the chapter.

<i>Class-Interval</i>	<i>Frequency</i>	
2.5-2.9 $\sigma$	0	
2.0-2.4 $\sigma$	1	
1.5-1.0 $\sigma$	4	
1.0-1.4 $\sigma$	10	
.5- .9 $\sigma$	21	
0- .4 $\sigma$	28	= ,I:P (Eq. 42, Ch. II)
-.5- -.1 $\sigma$	18	illustrating a frequency distribution
-1.0- -.6 $\sigma$	10	
-1.5- -1.1 $\sigma$	5	
-2.0- -1.6 $\sigma$	2	
-2.5- -2.1 $\sigma$	0	
-3.0- -2.6 $\sigma$	1	
	<u>100</u>	

in the social sciences, but hitherto statistics has lacked the notation for symbolizing the *distribution*, apart from its indices, in a formula.

Another illustration of the colon symbol which has proved the most useful symbol in all the S-theory is a set of time curves, such as the freight tonnage planned and achieved under the first Russian Five-Year Plan (first three columns of S. 31, Ch. II):

<i>Freight in Millions of Tons</i>			
	<i>Planned</i>	<i>Fulfillment</i>	
1929	165.0	187.6	
1930	185.0	238.7	
1931	210.0	258.3	= ,T <sup>-1</sup> :(I) <sub>i</sub> (Eq. 43, Ch. II)
1932	240.3	267.9	illustrating time series

In this matrix formula,  $\iota$  represents the four periods, each of one year (T), each with a corresponding aggregation of two ( $i = 2$ ) indices (I) of annual tonnage.

The colon is seen from these illustrations to assert that the entity or entities preceding it are each subclassified into, or have a corresponding subaggregation of, further entities as specified by the descripts following the colon. The colon asserts observed association of dependent entities on independent entities without necessarily implying a causal, probable, or sequential relation.<sup>22</sup> \* Its use is illustrated by almost every S-situation appended to the chapters of this volume.

\* For Eq. 44, Ch. II, see notes at end of the chapter.

## 2. THE CROSS-CLASSIFICATION SIGN (::)

Whenever the data are cross-classified rather than subclassified, the double colon will symbolize this. In a cross-classification neither variable is dominant or dependent; both are independent. The classification is reversible in that either variable can be considered as subclassified by the other. A simple and important example is a correlation scattergram or frequency surface. Thus the cross-classification of countries by Olympic athletic scores and by mean temperature in S. 7, Ch. II has the descriptive formula:

$${}_iI :: {}_jI : \underline{P}_p \quad (\text{Eq. 45, Ch. II})$$

illustrating the double  
colon sign of cross-  
classifying

This formula is statable in words as that for each of the class-intervals,  ${}_iI$  in number, of one indicator cross-classified with (::) each of the class-intervals,  ${}_jI$  in number, of a second indicator, there are corresponding (:) lists of plurels,  $\underline{P}_p$  in number, whose population is not stated ( $\underline{P}$ ).

Cross-classifying is an intermediate operational step between distributing characteristics separately and correlating them. (See also S. 12, 25, Ch. II; S. 1, 2, 3, 5, Ch. VI, and S. 1, Ch. XII.)

## 3. THE CORRELATION SIGN (•)

From a cross-classification such as a contingency table or a correlation scattergram, the tendency in that situation can be summarized in a single index such as the correlation coefficient. This coefficient is a mean of the products of the two indices measuring the characteristics (in sigma units). In vector algebra such a product is called the scalar product and is symbolized by a heavy dot between the two factors, i.e., the two indices correlated. As this dot is also the conventional statistical notation in the multiple correlation coefficient, it is adopted in S-theory as the sign of correlation. Thus the correlation of temperature and athletic prowess (S. 7, Ch. II) is .65 and is stated with algebraic generality in the type formula:

$${}_sI, \bullet {}_sI_{,,} = r, ,, \quad (\text{Eq. 46, Ch. II})^{23*}$$

illustrating the heavy  
dot as sign of  
correlation

\* For Eq. 47, Ch. II, see notes at end of the chapter.

Intercorrelation of variables, static or dynamic, is perhaps the most powerful quantitative tool the social scientists have yet developed for analyzing, measuring, and recording the inter-working *patterns* of societal phenomena.

#### 4. THE IDENTIFICATION SIGN

A fourth basic operator in S-theory is the prime. When a script is marked with a prime, or is replaced by the prime for brevity, it means that the script is singular and identifies one particular entity. For example,

${}^{\prime}P_p = {}^{\prime}P$  = one particular person in a particular plurel  
(Eq. 48a, Ch. II)  
illustrating the prime script

${}^{\prime}T = {}^{\prime}T$  = one specified date in one particular period  
(Eq. 48b, Ch. II)

${}^{\prime}L_i^{\dagger} = {}^{\prime}L_i^{\dagger}$  = one particular point in one specified area  
(Eq. 48c, Ch. II)

(I') = one particular index, a homosectoral index either T, I, L,  
or P  
(Eq. 48d, Ch. II) <sup>24</sup>\*

#### E. Summary of S-theory

The foregoing exposition of the indices, scripts, and operators<sup>\*</sup> of S-theory may now be summarized in a verbal statement and in a generalized algebraic formula, with more rigor than in the simple generalization that "people, environments and their characteristics may change."

"1. Every quantitatively recorded societal situation can be expressed as a combination of indices of *time* (T), of *space* (L), of *population* (P), and of the many *characteristics* (I) of people or their environment.

"2. Indices from these four sectors of societal data are modified by four scripts, namely, the usual mathematical *exponent* ( $^{\circ}$ ) and three other scripts specifying the number and nature of their *classes* ( $|_s$ ), *class-intervals* ( $_s|$ ), and *cases* ( $^s|$ ).

"3. These indices and scripts are operationally combined by

\* For Eq. 49, Ch. II, see notes at end of the chapter.

the usual signs for *adding* (+), *subtracting* (−), *multiplying* (×), and *dividing* (/), plus the signs for *aggregating* (:), *cross-classifying* (::), *correlating* (•), and *identifying* (') particular entities."

This verbal statement is symbolically expressed in the algebraic formula:

$$S = {}^s(T;I;L;P) {}^s \quad (\text{Eq. 50, Ch. II})^{25}$$

algebraic statement  
of S-theory

where the semicolon denotes any of the above eight operational signs for combining the indices, i.e., ( $; = + - \times / : :: \bullet '$ ).

This Equation 50 is the answer to the question which provoked this volume, namely, "Can a quantitative systematics of Sociology be constructed at present?" It is the general formula specifying the system of which all the seven hundred or more statistico-sociological concepts and relations, defined in this study by equations, are particular parts. *Every formula and equation in this volume can be shown to be a special case of or a derivation from, the basic Equation 50 specifying the S-theory in algebraic symbols. Furthermore, every statistical tabulation, graph, map, formula, or other quantitatively expressed set of data in the social sciences, can, if this S-theory is true, be expressed as a special case of Equation 50.* This is a very large claim. The evidence for it will occupy the rest of this book.

This matrix Equation 50, like most of the S-equations in this volume, is a descriptive more than a calculative equation. The difference is important. In a descriptive equation one member is observed and defines the other member which is set equal to it. Thus here, the left-hand member, S, is not directly observed independently of the right-hand member. In a calculative equation both members may be independently observable. If a term is unknown in the calculative equation, it can be solved for. As many unknowns can be solved for as there are independent simultaneous calculative equations. But in a descriptive equation the term defined is the unknown, and a second unknown in the same equation, taken alone, cannot be solved for. Defining concepts by descriptive equations based on objectively observable quantities is the first step, however, in manipulating them and discovering and testing relationships among the concepts. Such relationships are the most exact forms of that "understanding

of phenomena enabling prediction and control," which it is the function of science to achieve.

## II. UTILITY OF S-THEORY

No sensible social scientist is likely to read this heavy scientific exposition, weighted with algebraic symbols, unless he believes that it may be worth his while to do so. Consequently, before exploring the theory further in detail, and before marshaling the evidence as to the truth of the theory, some points as to its possible utility should be noted. Only if the theory bears promise of usefulness as well as truth is the reader likely to make the mental exertion of learning its new symbolic language. In discussing its utility, however, immediate application to remake society will not even be attempted. Rather, its uses in pure science will be sketched for making a better science which eventually is likely to be useful to society in a more fundamental, though less immediate, way. In short, these utilities are of more interest to the sociologist than to the social worker, of more interest to the "pure" than to the "applied" scientist.

The utility of S-theory may be conveniently summed up under the four headings of parsimony, precision, classification, and fruitfulness.

### A. Parsimony

Parsimony in the number of assumptions or concepts needed in a theory is a fundamental desideratum in science. Since the human span of attention is very limited, understanding the myriad phenomena about us progresses in proportion as we reduce them to fewer general classes, or actional uniformities called laws. There are twelve basic concepts peculiar to S-theory, namely, the four indices, the four scripts, and the four operational relations ( $TILP_{is}^{as} :: \bullet'$ ). Another dozen of their commoner compounds and a few less used symbols,  $((I)SVJ^{az\sigma m} | \underline{?};)$ , together with a dozen of the usual symbols of mathematics ( $\Sigma \pm \times / = > < \neq \equiv || ||$ ), suffice for writing the formulae describing all quantified societal situations. Thus, in the thirty-five graphs of this chapter there are thirty-six terms for various classes of population, which are all subsumable under our class script on P, and labeled "plurels" for short. In these graphs

there are seventy-four terms used for indicators of various characteristics which are all subsumable under I. In a sample of five hundred graphs, an estimate has shown that some fifteen hundred different terms can be replaced for systematic purposes by the eight homosectoral indices and scripts, and their combinations. Of course, no absurd claim is made that the fifteen hundred terms can be thrown away and replaced by eight for all purposes. Just as a mean usefully represents many individuals and is useful for dealing with a group, although inaccurate when dealing with the individual, so a general class concept is parsimonious only in those respects in which it can be substituted for the members of the class.

The sixty-one concepts recommended by a committee of the American Sociological Society as standard content for the introductory college course in Sociology can all be expressed in the twelve basic S-symbols, and a majority of them are thereby redefined in terms which can be measured (see Chapter XII for the list of these).

The use of algebraic symbols instead of words is a further parsimony. A symbolic equation is a more compact statement than a verbal sentence. For example, the statement of S-theory in the equation  $S = {}_s(T;I;L;P)_s$  takes a whole paragraph to express in words.

### *B. Precision*

A second utility of the theory is the emphasis it puts upon increasing the *precision* of observing societal phenomena.

For precision, objectivity is essential. Observations that vary subjectively from observer to observer may be satisfying to each observer alone, but have no precision for science which connotes agreed-upon knowledge verifiable by all competent observers. Science has found that the language of mathematics is the most objective language. When phenomena and their operational relations are expressed in mathematical symbols, subjectivity is reduced and precision is increased. The objectivity of S-symbols has been measured in percentage terms by means of a controlled experiment, which is described below.

Precision is further emphasized by the scripts which denote the degree of precision or limits of it. Thus, for example, the

class-interval script states the units of the index; the case script is much used to set the initial and terminal limits or range within which the formula holds; the case script often denotes ordinal quantities vs. the class script, which denotes cardinal quantities; the exponent of zero denotes primitive observation of qualities only; the exponent of one denotes more precise observation differentiating a quality into degrees of it; and the exponent of two denotes still more penetrating observation of the relations between quantified qualities. To write the formula of a situation requires precise thinking by the analyst and is a clear-cut technic for revealing lack of precision in the captions or presentation of the graph, table, or other data that are being analyzed.<sup>26</sup>

### *C. Classification*

A third utility lies in all that is connoted by the concept of *classification*—the arrangement of parts into an orderly whole, the relating of facts into an integrated system of knowledge called a science. The quantic table provides such a classification scheme, or frame of reference, for ordering societal phenomena, beginning with quantified phenomena and progressively extending on into the unquantified realm. The fact that every quantified societal situation without exception, among the fifteen hundred investigated, neatly falls into its class in this table demonstrates this classificational use. The absence of any standardized classification of the field of Sociology has been shown by Eubank's analysis of tables of contents of leading textbooks (Ref. 25, pp. 43 ff.). The system offered in this volume may prove a sterile flower destined to fade and be deservedly forgotten. But in presenting the S-theory as a candidate for wider standardization, the claim is made, to be judged by the reader, that no other sociological classification obeys the canons of classification so rigorously. These canons are that:

1. A classification must include *all* phenomena of the field classified—here the quantified segment (at present) of societal phenomena.
2. The classes must be mutually exclusive, i.e., the boundaries must be definite. The extent to which the rulings in the later chapters clearly differentiate the sectors and the different exponents of each is experimentally determined and measured.

3. There must be a uniform basis or principle which determines all the classes. The quantic table is essentially a cross-qualification on the basis of two principles taken conjointly—the principle of sectors and the principle of exponents. In a sense, developed in later chapters, the sectors may be viewed as the chief elemental classes of *content* of human knowledge about society—time, space, population, and the residue of all characteristics (whose further subdivision marks off particular social sciences). The exponents may be viewed as the chief elementary classes of the *form* of presenting that knowledge—the operational degree of superficiality-to-profoundness with which the data are presented.

The excellence of the classification of societal data given by S-theory is further suggested by the fact that the theory can be expressed in geometric terms as sheaves of vectors in an S-dimensional societal space. This shows that its categories are internally self-consistent.

#### *D. Fruitfulness*

A final utility of any scientific theory should be its *fruitfulness* for further research and the extension of knowledge. A good theory should be but the starting point of many further applications and explorations. At many points in the later chapters and in the comments on the graphs leads for further study or application will be pointed out. An explicit summary of the chief of these are listed as possible thesis topics for graduate students in Appendix IV pointing out some of the possibilities for further development of the field of Sociology as seen by the author. Briefly, such leads include: the extension of mathematical tools to qualitative phenomena opened up by our attribute hypothesis; the fuller use of matrix and vectorial algebra on societal data; the exploration of the quantic formula on data and in cells of the table hitherto undiscovered, whose properties, however, can be predicted in advance; the testing of the hypothesis of a maximum natural range of about  $12.5\sigma$  for human abilities and its application to distributions of cultural phenomena such as wealth; the trial of an equilibrium theory of societal action measuring the chief societal processes current among sociologists; further experimental and mathematical research on the hypothesis of epsilon elements providing equal and interchangeable “atoms” for

societal data; the development of the interrelation matrix as a tool for exact analysis of interhuman relationships, i.e., converting attributes to indicants and intercorrelating them into systems of increasing comprehensiveness.

### III. VERIFICATION OF S-THEORY

#### A. Criteria for Verification

The initial problem of the research reported in this volume was to construct a systematics of quantitative Sociology. Relevant data in graphic form was gathered, and a systematic "S-theory" was inductively built up. The last step of the scientific method calls for verification. Is the theory true?

To test truth the philosophy of science offers two general criteria, those of (a) *internal consistency*, and (b) *external correspondence*. Is the theory proposed logically consistent within itself and with the current body of theory that constitutes that science at a given date? Further, does the theory fit all the observed facts, i.e., does it correspond to external reality?

Supplementing these two principles of consistency and correspondence are two corollaries. These are the practical subcriteria of *agreement* and *utility*. What percentage of competent observers agree as to its being consistent with existing knowledge and corresponding to all the observed facts? If all agree, it is accepted as proven theory, otherwise it is still on trial to the extent of the disagreement. . . . Finally, even though accepted, it may be a sterile generalization. What is its utility for science? What further problems has it suggested in investigatable form? What further collecting of data or experimenting has it stimulated? What further generalizations, more inclusive, or better established, have grown out of it? Can it be applied to predict or control phenomena in ways which man values? These are the pragmatic subcriteria of utility.

#### B. The Criterion of Internal Consistency

The first criterion of truth, that of internal consistency, was applied by means of three technics, namely: (1) a *reliability* technic consisting of independent analyses by two different persons; (2) a *geometric* technic consisting in translating S-theory

into terms of vectors with their points, lines, and angles; (3) a *paraphrasing* technic consisting in translating the concepts of current sociological theory into S-formulae to determine the extent of the consistency of S-theory with the existing state of the science.

## 1. RELIABILITY OF S-THEORY

Inconsistencies in interpreting the definitions of the concepts and the notation are experimentally revealed when two persons write the formulae for S-situations independently and then note their discrepancies and study the reasons for them. This was done for a sample of five hundred graphs (S-situations) from our collection culled from the postwar literature of all the social sciences.<sup>27</sup>

Counting as one discrepancy each unit of difference between two analysts in any digit of the quantic number, a percentage of discrepancies was calculated. The percentage of discrepancy between the *quantic formulae* of the two analysts was found to be three percent. This three percent of discrepancy means ninety-seven percent of agreement between the two analyses. *The reliability of the quantic classification of S-theory is thus shown, within the limits of this experiment, to be ninety-seven percent ( $\pm 76\%$ , i.e., with a standard error of sampling of less than one percent).*

A second experiment was made by Mr. Bruce Billings (Harvard, M.A. in Physics) whose mathematical training was far superior to Miss Jurdak's. He received no personal instruction from the author and secured his knowledge of S-theory solely from study of the first manuscript of this book. On a smaller sample of fifty-five graphs which he had not previously inspected, culled from various chapters of this book, his independent writing of the quantic formula in November 1938 showed ninety-nine percent ( $\pm 45\%$ ) of agreement with the standard version published here. The reliability of the quantic classification by this experiment was then ninety-nine percent.<sup>28</sup>

The discrepancy index of reliability was also calculated for the full descriptive formulae of the five hundred situations analyzed by Miss Jurdak. Counting each letter or symbol (including their positions) in which the two independent analyses differed from

each other as one point of discrepancy, the percentage of discrepancies was calculated. The percentage of discrepancy was found to be seven percent, or ninety-three percent of agreement. The reliability of the *descriptive formula* is then, as far as this experiment goes, ninety-three percent ( $\pm 1.14\%$ ). This reliability may be analyzed by subsamples as follows:

<i>Sample of S-situations</i>	<i>Average Percent of Agreement</i>
1st 100	89%
2d 100	90%
3d 100	92%
4th 100	95%
5th 100	97%
All 500	93%

One further analysis was for one person to rewrite the formulae after a month. The agreement between the first and second writings was ninety-eight and five-tenths percent for the quantific formulae and ninety-six and six-tenths percent for the descriptive formulae. These percentages of discrepancy show the extent to which the rules for converting sets of data into S-formulae (see Appendix II) are inadequate and need further refinement towards the goal of no discrepancies.

Whether this is an excellently high reliability for this kind of data is unknown for lack of similar experiments to furnish a norm. For, as far as the author knows, there has never been a similar controlled experiment to measure the objectivity of any of the many classifications proposed by social scientists for their science. Classification construction has been an art, practiced by the writers of textbooks and systematic works without any attempt hitherto to reduce their classifications as hypotheses to the scientific test of how consistently it gives similar results in the hands of different users. Yet to appraise the excellence of any classification, its reliability is the first thing to be determined and reported. Now that a technic has been developed, scientific standards should require that classifications in the future have their reliability quantitatively determined.

The important point about this observed reliability of ninety-seven percent is that it can be increased with the practice and training of the analyst. In this experiment the second analyst,

Miss Jurdak, had had only a one-semester course in Statistics and no Co-ordinate Geometry or Advanced Algebra as a background. But the percentages of agreement showed a trend to rise with practice.<sup>29</sup> This is to be expected since every discrepancy reveals an ambiguity in the Rules which can be recorded toward preventing the recurrence of that kind of discrepancy.

## 2. GEOMETRY OF S-THEORY

The second technic for determining the internal consistency of S-theory was to convert it into geometric terms and see whether a spatial model could be constructed. Such a model is diagrammed in S. 35, Ch. II.

Every index is represented as a vector which is a directed line, i.e., a line with length and an angle with another reference line. Case scripts state points on its vector, class-interval scripts specify the line-sects subdividing the length of the vector, and class scripts specify the number of different vectors in a sector. With all vectors starting from a common origin of zero they spring out as a sheaf of lines into a Euclidian space of  $n$  dimensions,  $n$  being the total number of indices in the situation as recorded. A zero exponent denotes a point or a dimensionless space; an exponent of one denotes an extension, a vector; an exponent of plus two denotes an angle between two vectors and also the area of the parallelogram determined by the two vectors. (See S. 12 and 13, Ch. VI for exact formulae.) A correlation coefficient (symbolized by the heavy dot) is the cosine of the angle between the two vectors representing the two indices correlated. A table of inter-correlation coefficients, considered as an algebraic determinant, specifies all the angles between every pair of vectors in the situation. An index that is subclassified by the colon symbol (as Eq. 39, Ch. II) is a vector, which is like a centroid and splits up into a sheaf of vectors, each of which represents one of the sub-classes.

In this scheme qualities are geometrically represented by directions, and quantities by lengths of lines. Correlations measure degrees of angle and, therefore, quantify the degree of qualitative difference between two entities.

Vectors can be added, subtracted, and multiplied, by the rules of vectorial algebra. They can be transposed into other sets of

vectors, such as a system of rectangular (Cartesian) co-ordinates. The dimensions of a societal situation are the number of its vectors. Under certain conditions (stated in Ch. VI) the vectors of a situation can be re-expressed in terms of fewer vectors, i.e., in a space of fewer dimensions. This geometric interpretation is expanded in Chapters III and VI. Here it may be epitomized in the geometric formula, recalling that overlining denotes vectors:

$$\begin{aligned} \text{number of points} &= \overline{i}^i \text{ (deals with directions)} \\ &\quad (I) = \text{scalar length} \\ \text{number of sects} &= i^i = \text{number of vectors} \end{aligned}$$

(Eq. 51, Ch. II)

the geometric interpretation of  
S-theory

In this geometric interpretation of S-theory, every S-situation is represented by an n-dimensional hyperspace which may be referred to as "S-space." There are usually as many dimensions as denoted by the class scripts, which are weighted in each sector by the exponent as specified more precisely in later chapters:

$$n = |_t + |_i + |_s \times |_1 + |_p \quad (\text{Eq. 52, Ch. II})$$

the number of dimensions in a societal situation

This S-space is divided into four subspaces called sectors—those of the temporal, indicatory, spatial, and populational types of vectors. Each of these sectors is subdivided into as many subspaces (which may be called vectorial subspaces) as there are vectors of that type as denoted by the class script,  $|_s$ . Each vectorial subspace may be further subdivided into normals, which are a special type of mutually perpendicular co-ordinates, the number of which is denoted by the exponent. Each vector is subdivided into sects as denoted by the class script. Finally the sects, or the vectors directly, have specified points on them as denoted by the case script. In S-notation (where the symbol:

is read as, "is subclassified into") this is expressed in the formula:

Verbal formula for S-space:	S-space	:	sectors	:	vectors	:	normals	:	sects	:	points
Symbolic formula	$\bar{S}$	:	$\bar{I}'$	:	$\bar{I}$	:	$\bar{I}^t$	:	$\bar{I}^i$	:	$\bar{I}^p$
Sociological equivalents:	Quantified societal Situations	are analyzed into	4 Types of Indices	each type having	a specified Number of Indices	each operationally developed according to	the Exponent	and subdivided into	a specified number of Class-Intervals	and subdivided further into	a specified number of Cases

A further geometric model of the quantic formula alone,  $S^s = (T^t; I^i; L^l; P^p)$ , has the four sectoral exponents as four variables. It can be best represented in a four-dimensional space. As this cannot be visualized, S. 33, Ch. II presents a three-dimensional model with a fourth dimension (of length) suggested as a subclassification (of time). This is an expansion of the Quantic Table above. The usual exponents of time,  $T^0$ ,  $T^{\pm 1}$ ,  $T^{\pm 2}$ , define three "avenues" in this model of blocks; the more usual exponents of indicators define three "streets"; the exponents of population define three "stories." The bottom left-foremost corner represents the simplest situation with a quantic number of 0;0;0;0. There is a steady increase in the exponents and the number of sectors involved (with non-nullary indices) in zones moving from that corner to the top right-rear corner whose quantic number is 8;2;0;2 (or 8;2;3;2, if length were represented properly).<sup>30</sup>

These geometric models of S-theory serve as indirect evidence of its internal consistency, for geometric constructs of lines and angles will often bring inconsistencies out into high relief.

### 3. PARAPHRASING OF CURRENT CONCEPTS IN S-FORMULAE

The third technic for testing internal consistency of a theory is to compare it with pre-existing theory. A new theory which harmonizes with the existing body of theory in a science is to be preferred, other considerations being equal, to a new theory which clashes with previous theory. Towards this aim the existing current concepts of Sociology have been translated into S-formulae

throughout this volume. The summary of such paraphrasing, or translating, is presented in Chapter XII with the finding that the majority of current concepts of the semi-officially endorsed list can be re-expressed in S-formulae, and many of them are thereby reduced to measurable terms. The minority of current concepts which are not reducible into S-formulae are so subjective and vague, that in the author's judgment, the prediction that many of them will in time be discarded by sociologists has high probability.

### *C. The Criterion of External Correspondence*

#### 1. THE REDUCTION OF QUANTIFIED SOCIETAL SITUATIONS TO S-FORMULAE WITHOUT EXCEPTION

To the basic question of correspondence of the theory to external reality a rigorously scientific answer has been attempted. This answer consists of the demand that *every single quantitatively recorded situation in the social sciences shall be describable in an S-formula without a single exception* or departure from the usual notational rules. Quantified situations in the form of tables, graphs, maps, pictures, formulae, and paragraphs were collected, sampling the fields of Sociology, Economics, Political Science, Anthropology, Education, Religion, and Philosophy. In order to reduce possible selective biases, *every* graph, table, etc., in a volume, if it was scanned at all, was included in the collection. Thus the seven chief sociological journals <sup>31</sup> were culled for every available issue of the past ten years. A dozen textbooks of Sociology, Statistics, and other social sciences, and other representative books, such as *Recent Social Trends*, were similarly harvested. Over fifteen hundred photostats were made of such situations, mounted, and presented to the author and his assistants to reduce to formula. The clerks collected mechanically under instructions not to omit one table, graph, map, or formula in the volumes scanned. The analysts then proceeded to analyze each photo-stated S-situation and write its formula, permitting no situation to remain unanalyzable. *Not one quantitatively recorded societal situation has yet been found which cannot be described by the S-theory formula.* (Eq. 50, Ch. II.)

Of course, the theory was not adequate at first and was con-

tinually modified to meet new situations.<sup>32</sup> In all probability it will be modified further in the future. Every situation resisting analysis became crucial and resulted in modifying the theory.<sup>33</sup>

A single unanalyzable quantitative situation would disprove the theory in the sense of demonstrating that it was not adequate, as it claims to be, to describe *every* such situation in terms of S-theory concepts in a pattern specified by the S-formula. Since indicators are defined as everything other than time, space, and population, it is logically inevitable that these four categories should include all societal data without any remainder. The real problem is not whether an S-formula can be written for every set of societal data, but whether different analysts will write the same formula. This is the question of reliability answered above.

The graphs illustrating the theory in this volume are, of course, selected ones. The collection from which they were selected is a sampling of the literature. But this sampling was done on an objective basis regardless of whether or not these graphs might illustrate S-theory, and is not a selection of illustrations gathered to support a theory with non-supporting illustrations ignored. This inductive approach of insisting that the theory fit *every* recorded quantitative societal situation encountered is what is meant by the statement above, that a rigorously scientific answer was attempted to the question as to how well the theory corresponds to external reality.

#### *D. The Subcriterion of Agreement of Competent Persons*

The argument for the utility and for the truth of S-theory has been stated up to this point by one person. There remains the evaluation of other sociologists and social scientists for applying the corollary subcriterion of the consensus of opinion of competent students.

The opinion of other scientists and sociologists, endorsing, qualifying, amending, or rejecting the S-theory is not available in advance of its publication. Note may be made, however, of a few indications of appraisal of parts of the manuscript.

The theory for measuring societal forces, developed in Chapter XI was first published in Refs. 12 and 18, and its standard error in Ref. 16. This theory was reviewed or commented upon

favorably by G. A. Lundberg in *Social Forces* (Ref. 44);<sup>34</sup> by F. S. Chapin in his presidential address to the American Sociological Society (Ref. 7, pp. 6, 9)<sup>35</sup> and in his book, *Contemporary American Institutions* (Ref. 6, p. 347);<sup>36</sup> and by H. A. Phelps in his book, *Principle and Laws of Sociology* (Ref. 55, pp. xii, 45, 51, 130, 160).<sup>37</sup>

An article embodying the section of the tension theory treated in Chapters V and X was rejected by the *American Journal of Sociology*, and was later solicited by the editor of the *American Sociological Review* and published in the February, 1939, issue (Ref. 19). H. A. Phelps produced part of an early draft of it in his book (from a typescript in clumsy notation which had been shown to him by the author). This draft interested Lundberg, then chairman of the Research Section of the American Sociological Society and in charge of its annual canvass of research in progress, to the extent of proposing collaboration in preparing one of two volumes on systematic quantitative sociology. As this work developed he came to Syria in the summer of 1937 to study S-theory and plan what developed into his companion volume, *Foundations of Sociology* (Ref. 43), and the present volume. His encouragement and criticism have been the two most important influences on the author in developing and publishing S-theory.

In order to get criticisms of S-theory, copies of a six-page abstract were circulated among a group of sociological, psychological, and statistical colleagues in England and the United States. From some of these, critical comments have been received on this abstract or on other isolated parts of the manuscript: Read Bain (Miami), Arthur F. Bentley (Indiana), Jessie Bernard (Washington), Bruce Billings (Harvard), F. Stuart Chapin (Minnesota), Irvin Child (Yale), Michael Choukas (Dartmouth), Earle E. Eubank (Cincinnati), Karl Holzinger (Chicago), Clark Hull (Yale), L. L. Thurstone (Chicago), Charles Malik (Harvard), graduate students of H. A. Phelps (Pittsburgh), and Nicholas Reshevsky (Chicago). The assistance of their comments is gratefully acknowledged. They are in no way responsible for whatever mistakes the author may have made in the present volume.

The entire text has been critically read by Mr. Billings to check mathematical formulae, notation, and references, and by Mr. Lundberg to check underlying assumptions, and the author

has attempted to incorporate fully their wealth of critical notes and corrections in the text.

A resumé of S-theory appeared six months after this volume had gone to press in the *American Sociological Review* in October 1939 under the title "An Operationally Defined System of Concepts for Sociology" (Ref. 19a).

#### IV. SAMPLE SITUATIONS REDUCED TO S-FORMULAE

A selection of thirty-five descriptive S-equations will be presented next, together with a few calculative ones. They are taken from all the quantic topics of the succeeding chapters to give a fuller view of the range and forms of the S-theory throughout the social sciences, before entering into its detailed exploration one quantic topic at a time. The equations are grouped by the steps of the scientific method, as outlined in the preceding chapter, to illustrate methodological uses.

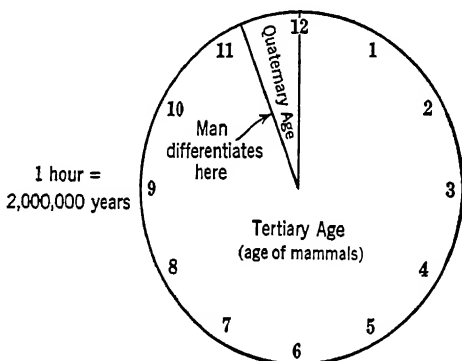
Each graph bears its quantic number, [<sup>s</sup>, classifying it into a cell of the Quantic Table and showing to which of the subsequent chapters it belongs. After the full descriptive formula is written the "legend" specifying what each index and script represents in the given situation. The S-symbols analyze the content of societal situations, aside from their relationships, into eight standardized general concepts—the four sectors and their four scripts. The legend states in words the particular species or form which each of these eight general concepts represents in the specific situation. The descriptive formula classifies the constituents of the situation into genera; the legend subclassifies each genus into species.

The legend may be read horizontally or vertically. Horizontally each symbol is defined in the phrase aligned with it. These definitions have connective phrases in intermediate rows so that the right-hand members by themselves may be read vertically from top to bottom giving a connected description of the situation in terms of the concepts of S-theory.

After the first thirteen S-situations, illustrating the four sectors and their commoner exponents, the remaining twenty-one illustrate the broad steps of the scientific method in: (1) observing and recording facts, (2) systematizing them, (3) inducing principles, and (4) verifying them.<sup>38</sup> These thirty-five situations, each described by the appropriate form of the S-theory formula (Eq. 50,

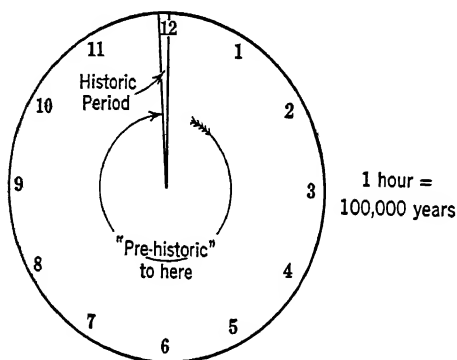
Ch. II) constitute a preliminary sample of the way S-theory corresponds to external reality. More detailed evidence is presented in the three hundred situations appended to the chapters of this volume. These are a sample of the facts to which any systematic sociological theory must be fitted.

## S. 1



## THE ORIGIN OF MAN

In this diagram the author follows the more liberal geological estimate of about 25,000,000 years since the beginning of the Tertiary. Man differentiated about half-past eleven.



## THE QUATERNARY AGE: "PREHISTORIC" AND HISTORIC TIME

This diagram is designed to indicate the utter chronological insignificance of the period of recorded history when compared with the whole period of human existence on the earth.

Descriptive formula:  $S_1 = t : {}^a_u T^{+1}$

Quantic number = 1;0;0;0

Legend:

$S_1$  = The situation

into

records

$u|$  = 2 periods (prehistoric and historic)

$T$  = time

with

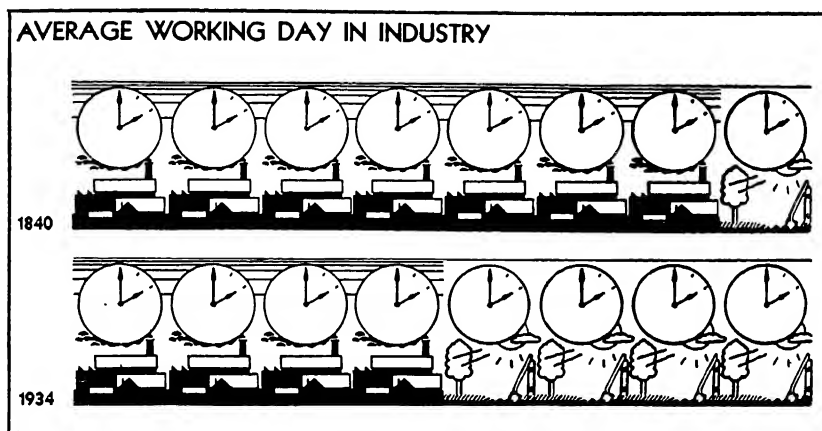
divided into

$t|$  = 2 ages (tertiary and quaternary)

$a|$  = initial dates stated (since the periods are unequal)

$:$  = and subdivided

## S. 2



Ref.: Research Bulletin, National Educational Association, Vol. XII, No. 5, November, 1934.

Descriptive formula:  $S_2 = \underline{P}_, : {}^t T^{-1} : T^{+1}_,$

Quantic number = 91;0;0;1

Legend:

$S_2$  = The situation

$T^{-1}$  = a velocity of change

is a record of

$_,$  = classified in year units

$\underline{P}_,$  = an unnamed plurel

in

between

$T^{+1}$  = duration of the working day

$^t|$  = 2 dates (1840 and 1934)

$_,,$  = classified in hour units

stating

Comment:

The average velocity of the process of shortening the working day is almost four minutes per year in this population and period (360 minutes/94 years = 3.8 minutes). In relative terms this is a decrease of 45% per century.

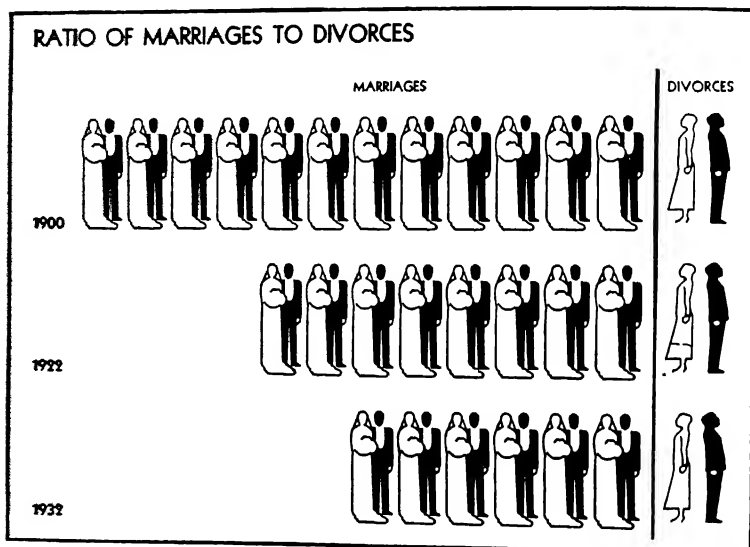
*Comment on notation:*

1. The two particular classifications of time (into years and hours) are two time indices.

2. In terms of matrices, the data can be arranged in a  $2 \times 1$  matrix as denoted by the multiple date script of two,  $\uparrow$ , and the corresponding singular script on  $T, \cdot$ . Thus, there are but two cells in this matrix, a duration stated for each of the two dates.

3. The combination of a positive and a negative exponent of unity is denoted in the quantic number,  $|$ , by the digits 91.

## S. 3



*Ref.: Research Bulletin, National Education Association, Vol. XII, No. 5, November, 1934, p. 254.*

*Descriptive formula:*  $S_3 = \uparrow T^{-1} : (\underline{P}, \underline{P}, \uparrow T^{-1})$

*Quantic number* = 8;0;0;19

*Legend:*

$S_3$  = The situation records

$\uparrow$  = 3 periods

of

$T$  = a year

$\uparrow$  = (with corresponding dates stated)

$:$  = with a corresponding

$()$  = index

for each, composed of

$\underline{P}$ , = the unstated number of persons marrying

divided by

$\underline{P}, \uparrow$  = the unstated number of persons divorcing

$T^{-1}$  = per year

*Comment:*

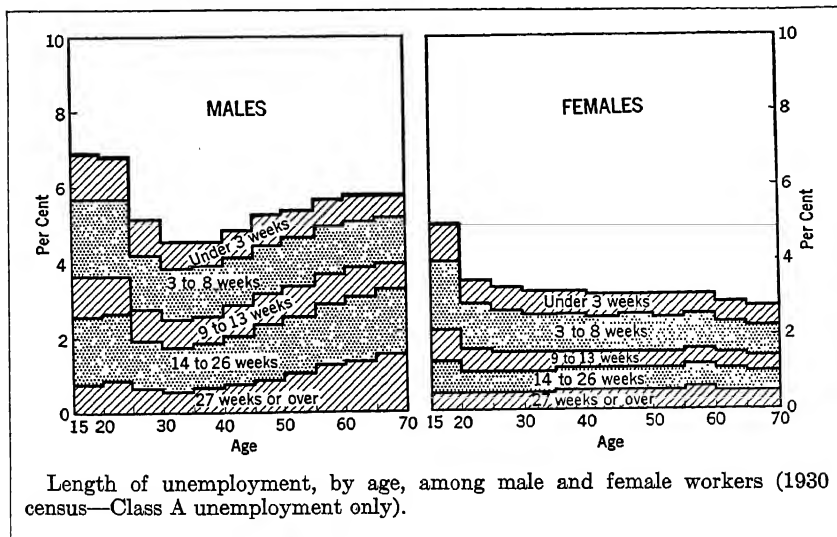
This acceleration of the divorcing process can also be expressed as a qualitative societal force as defined in S. 21, Ch. II: the amount of an acceleration of a change

times the population changed,  $F = T^{-2}I^0P^{+1}$ . The attribute, divorce, which is multiplied by the "pure" P can be written explicitly. The existence or non-existence of a divorce becomes an all-or-none type of change represented by the attribute with its value of unity, when present, and zero, when absent,  $^{+1}I^0$ . This index, F, measures the net societal force accelerating divorcing, in spite of whatever resisting forces may be operating. It is .0026 units of force expressed relative to units of marriage, i.e., .0026 new divorcees per marrying person per year per year, in the United States from 1900 to 1932.

*Comment on notation:*

1. The underlining denotes that the absolute number of that quantity is indefinite.
2. The use of parentheses denotes an index, (I); but here its factors are written explicitly.
3. Since the divorcing persons are not a part of the population getting married, two P's are written in a ratio ( $PP^{-1}$ ) instead of a " $\%P$ " meaning a population in percentage units.
4. The convenient use of the correspondence symbol, the colon, in the date script may be noted. Since the periods are intermittent ones, the date script specifies this fact, in asserting that initial ( $^a$ ) and corresponding terminal ( $^{:2}$ ) dates are recorded.
5. The acceleration is denoted by the exponent of  $-2$  in the time sector, which is represented by the digit 8 ( $=10 - 2$ ) in the quantic number in order to obviate minus signs.

S. 4



*Descriptive formula:*  $S_4 = {}^v({}_tT^{+1} :: {}_uT^{+1} : \%P)_p$

*Quantic number* = 2;0;0;1

*Legend:*

$S_4$  = The situation

records for

${}^v|$  = a particular date, 1930

for each of

$|_p$  = 2 sex plurals of class A workers  
in the U.S.A.

$|_p'$  = males,  $|_p''$  = females

a matrix in which

$|$  = 11 age periods

each of

$T$  = 5 years' duration

${}^v|$  = beginning at age 15

$::$  = are cross-classified

with

${}_uT$  = 5 unequal durations of unemployment

$:$  = with the corresponding cell  
entries of

$\%P$  = the percent of persons

$-$  = (absolute numbers not being  
stated)

*Comment:*

One trend in this depression was a great deal of unemployment amongst the youth who came newly onto the labor market.  ${}_uT : ({}_tT \cdot P, > 0 > {}_tT \cdot P_{,,})$ , which states, that for each unemployment duration,  ${}_uT$ , increasing age,  $({}_tT)$ , above age 30,  $({}^v|)$ , correlates positively with the number of men unemployed and correlates negatively from age 15 upwards,  $({}^v|)$ , with the number of women unemployed. Men's chances of employment went down with age, while women's chances went up slightly. This illustrates the use of S-notation to state trends and findings in quantitative data, without the necessity of further legend, other than the legend describing the whole situation.

*Comment on notation:*

1. The double colon denotes a cross-classification which has not been summarized in a single index, such as a correlation coefficient (or scalar product of two vectors).

2. Ten correlations are aggregated in the inequality above, since  ${}_u| = 5$  and  $|_p = 2$  and  ${}_u| \times |_p = 10$ .

3. That the five unemployment periods,  ${}_uT$ , are of unequal duration, is shown by the date script,  ${}^v|$ , symbolizing the initial dates of those periods. If the periods are equal, explicit assertion of the initial date of each *after* the first is unnecessary. Asserting initial dates implies unequal periods. (See Rule 53, Appendix II.)

## S. 5

### SOME OF OUR MORE IMPORTANT SCIENCES

#### I. THE ABSTRACT SCIENCES

- A. Mathematics
- B. Logic
- C. Metaphysics

#### II. THE PHYSICAL SCIENCES

- A. Chemistry
- B. Physics
- C. Astronomy
- D. Earth Sciences

#### III. THE BIOLOGICAL (LIFE) SCIENCES

- A. Botany (plant life)
- B. Zoölogy (animal life)
- C. Physiology (activities of organisms)
- D. Morphology (form and structure)
- E. Ecology (response to environment)
- F. Bacteriology
- G. Embryology

S. 5 (*Continued*)

## IV. PSYCHOLOGY

## V. THE SOCIAL SCIENCES

- A. History
- B. Economics
- C. Political Science
- D. Sociology
- E. Anthropology

## COURSES IN BOTANY AT THE UNIVERSITY OF CHICAGO

Elementary Botany	Ecological Anatomy
Elementary Plant Physiology	Experimental Ecology
Elementary Ecology	Experimental Field Ecology
Methods in Plant Histology	Geographic Botany
Organic Evolution	Physiographic Ecology
The Local Flora	Ecological Surveying
	Forest Ecology
General Morphology of Thallophytes	Field Ecology
General Morphology of Bryophytes and Pteridophytes	Applied Ecology
General Morphology of Spermatophytes	Plant Genetics
Special Morphology of Algae	Seminar in Evolution and Heredity
Special Morphology of Fungi	Seminar in History of Botany
Special Morphology of Bryophytes	Research in Morphology
Special Morphology of Pteridophytes	Seminar in Physiology
Special Morphology of Gymnosperms	Research in Physiology
Special Morphology of Angiosperms	Seminar in Ecology
General Morphology of Fossil Plants	Research in Ecology
Special Morphology of Fossil Plants	Research in Plant Genetics
Cytology	Research in Paleobotany
General Plant Physiology	Plant Chemicals
Plant Microchemistry	Plant Production in the United States
Plant Physics	Growth and Movement

The chart [above] shows something of the range of our modern sciences. It is a very abbreviated chart indeed, giving only main divisions. It is, even in its abbreviated form, a great view of the results of man's long development, and it is a great promise of the future, for never in the past has there been even a near approach to the powers which are in man's keeping, now that he has access to general laws of science.

The purpose of the list of courses [above] is to show how far we have gone in our study of one science. It is by no means an extreme case. Others have been carried farther. Do not try to understand all the terms.

Ref.: Marshall, Leon C., *The Story of Human Progress*, Macmillan, 1925, p. 157.

*Descriptive formula:*  $S_5 = I^0_{i,j,k}$

*Quantic number* = 0;0;0;0

*Legend:*

$S_5$  = The situation  
records

$I^0$  = a qualitative characteristic  
(science)

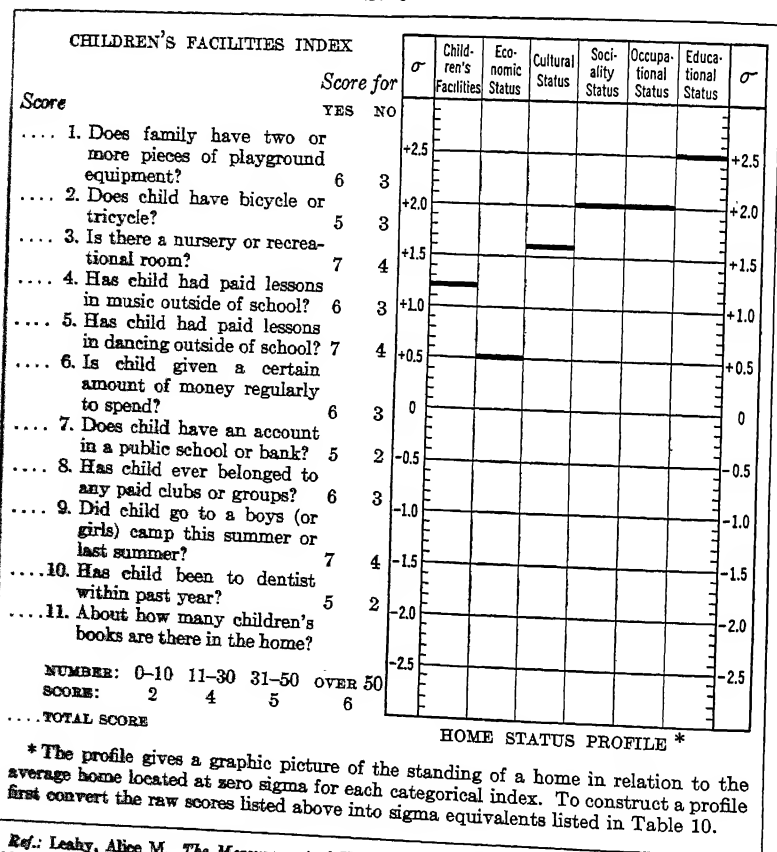
: = subclassified into

$|_i$  = 5 first-order classes  
 $|_j$  = 19 second-order classes  
and  
 $|_k$  = 43 third-order classes

*Comment:*

Note the function of the "legend." Only the four items below, which state the specific kinds or amounts in this situation, are essential.  $I^0$  = science,  $i = 5$ ,  $j = 19$ ,  $k = 43$ . The rest of the legend simply puts the formula into words for the benefit of the reader unpracticed in the notation. The S-formula, in general, states the pattern or structure of the situation in terms of about a dozen standard concepts (4 indices, 4 scripts, 4 chief operators). The legend states what particular forms or values these take in the given situation. A qualitative analysis in Chemistry giving the structure of elements, and a quantitative analysis of a chemical compound giving the amounts of each element, are somewhat analogous to the formula and legend of S-situations. The analogy is misleading, however, if the distinction is between quality and quantity, since the distinction in formula and legend is between general and specific. The descriptive formula analyzes the situation into a few general concepts, and the legend notes the specific kind of entity or amount of that entity, which is represented by each general concept.

## S. 6



*Descriptive formula:*  $S_6 = {}^L P : {}_i I_i$

*Quantic number* = 0;1;0;1

*Legend:*

$S_6$  = The situation

on

records for

${}_i |$  = a scale of 11 class-intervals

${}^L P$  = a particular person

for each of

- = who is not named (i.e., indefinite identity)

$|_i$  = 6 indicants of "home status"

$I$  = his I-units (i.e., scores in class-intervals of  $.5\sigma$ )

*Comment on notation:*

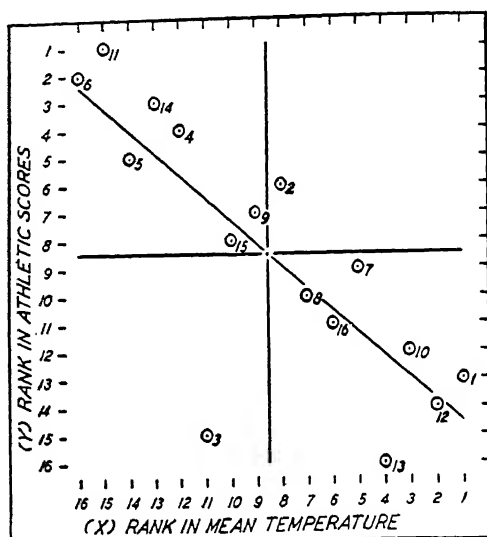
1. The use of the class-interval script to specify the unit used is here illustrated. Such symbols as,  ${}_s |$ ,  ${}_i |$ ,  ${}_6 |$ , state the kind of units or class-intervals, and not the number of them.

2. Since the value of  $P$  is unity, the colon could be omitted, as it makes little logical difference whether it reads "for one person there correspond six scores," or "the product of one person and each of his six scores." The connotations of a matrix at its lower limit merge into those of an "index" which is its limiting case.

3. The scripts show the  $S$  to be a second-degree matrix of order  $11 \times 6$ , as  ${}_i | = 11$ , and  $|_i = 6$ . The six scores of this person are a first-degree matrix,  $|_i = 6$ . The six scales constitute the second-degree matrix.

4. In this  $S$ , the only information given by the legend, which is not already explicitly stated by the formula, is, that there are eleven class-intervals, and that the indicants are six in number,  ${}_i |$ , and are of the kind named "home status." The legend could be condensed to  ${}_i | = 11$ ,  $|_i = 6$ ,  $I$  = "home status."

## S. 7



Correlated variability of specified countries in respect to athletic scores, in Olympic Games, 1920-24 (Y, or vertical scale), and mean temperature of country (X, or horizontal scale), showing a tendency in favor of the cooler climates. The athletic scores are ranked relative to population. The ranks of the countries are as follows (the rankings of each country are charted by the positions of the circled dots, numbered as below, and read on the two scales):

No.	Country	X	Y	No.	Country	X	Y
1.	Australia.....	1	13	9.	Holland.....	9	7
2.	Belgium.....	8	6	10.	Italy.....	3	12
3.	Czechoslovakia.....	11	15	11.	Norway.....	15	1
4.	Denmark.....	12	4	12.	South Africa.....	2	14
5.	Esthonia.....	14	5	13.	Spain.....	4	16
6.	Finland.....	16	2	14.	Sweden.....	13	3
7.	France.....	5	9	15.	Switzerland.....	10	8
8.	Great Britain.....	7	10	16.	United States.....	6	11

Ref.: Reinhardt, J. N. and Davies, G. R., *Principles and Methods of Sociology*, Prentice-Hall, 1932, p. 29.

Descriptive formula:  $S_7 = {}^{T^0} : {}^1I :: {}^1I : \underline{P}_p$

Quantic number = 0;2;0;1

Legend:

$S_7$  = The situation  
records for

$T^0$  = 1920 or 1924

${}^1I$  = the ranks of one indicant, temperature

:: = cross-classified

with

${}^1I$  = the ranks of another indicant, athletic scores,

: = with a corresponding

$\underline{P}_p$  = frequency of national plurals in each cell

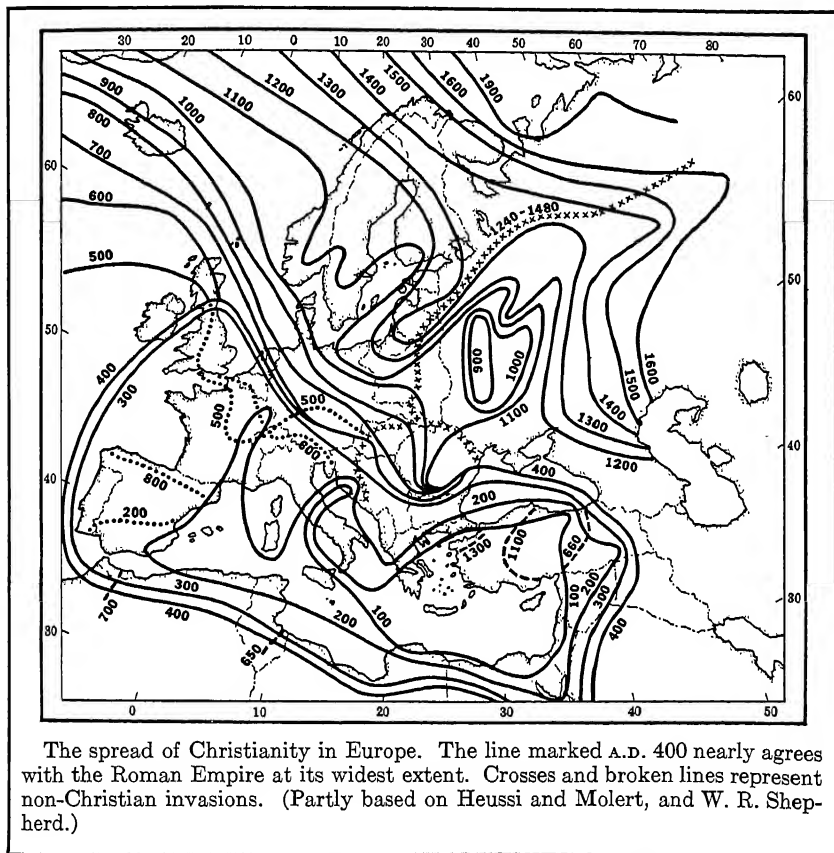
*Comment on notation:*

1. If the correlation coefficient were shown, summarizing the matrix in a single index, the dot operator would replace the double colon. The correlation is .65 ( $= \sigma I_{I,J}$ ).

2. Whenever the number of persons in each of the plurels is not stated, the P is underlined, denoting plurels each of indefinite size.

3. The case script denoting points rather than class-intervals is used here, as ranks signify ordinal points and not equal class-intervals.

## S. 8



Ref.: Taylor, Griffith, "Environment and Nation," *Amer. Jour. Soc.*, Univ. of Chicago Press, Vol. XL, No. 1, July, 1934, p. 28.

*Descriptive formula:*  $S_3 = \iota T^{-1} : \underline{L}_1^2,$

*Quantic number* = 9;0;2;0

*Legend:*

$S_3$  = The situation

of

records

$\underline{L}_1^2$  = the increase of Christianized area in Europe

$T^{-1}$  = the velocity (per 100 years)  
in each of

$|,$  = compared with zones of non-Christian invasion

$\iota|$  = 19 centuries

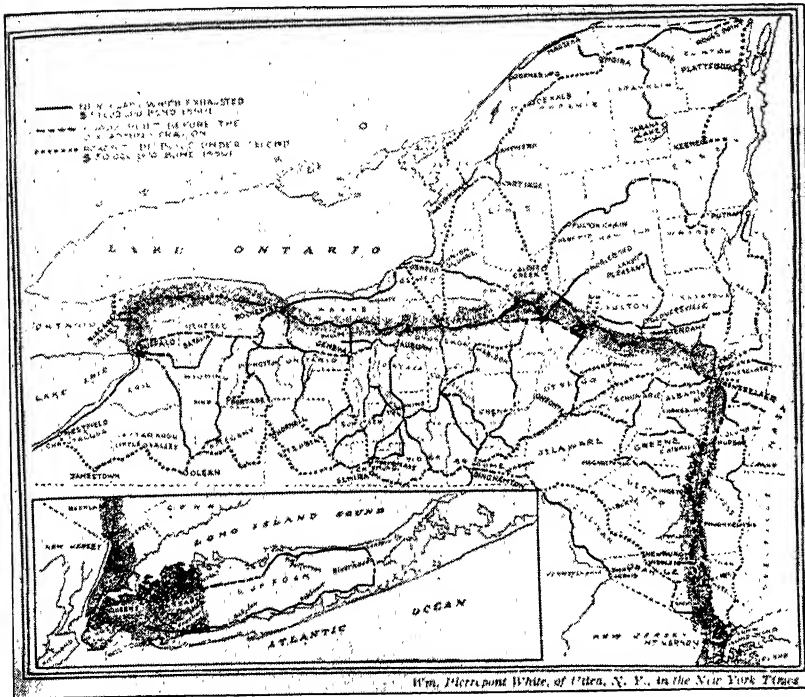
*Comment on notation:*

1. The diffusion of culture might be more explicitly stated by writing the formula,  $\iota T^{-1} : (I^0 L^2),$  in which the attribute,  $I^0$ , represents "Christianity." Whenever a quality characterizing an area is stated, the "pure" area has an attribute implicitly multiplied with it, and this may be written explicitly if desired. (Similar implicit, or "diphthonged," attributes usually characterize time and population which seldom occur "pure" in societal situations.) The "diphthong" is a product of a quality times a quantity.

2. Since the exact area of the zones is not stated in a map, the  $L^2$  is underlined, denoting indefiniteness.

3. Note that the zones of non-Christian invasion are compared, as set off by the comma, and denoted by a singular descript, since there is but one such zone for each of the 19 centuries. The effect of a multiple descript in an S-formula continues throughout all entities written to the right of it. There are thus  $\iota|$  zones.

## S. 9



Map Showing 3500 Miles of Completed and Proposed State Roads, in New York's Proposed 12,000-Mile System. The Shaded Portion Shows a Strip Ten Miles Wide which Contains 90 Per Cent of the Taxable Valuation and 80 Per Cent of the Population.

Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, McGraw-Hill Book Co., 1914, p. 222.

Descriptive formula:  $S_0 = {}^1L_{1:m}^2 : (L_n^+!_0, I, \%P)$       Quantic number = 0;1;21;1  
Legend:

$S_0$  = The situation

records

$L^2$  = a region, New York State,

- = (area not stated)

in

$|_1$  = 2 classes (counties and zoning)

with their

$|_m$  = subregions, i.e., counties, or 10-mile strip and non-strip areas

and

$^1|$  = specific points, namely cities,

and corresponding to these regions are

$L_n$  = lines in 2 classifications (roads and boundaries)

: = subclassified into

$|_0$  = 3 categories of roads

and also corresponding to the regions is

$I$  = an indicant of taxable valuation

and

$\%P$  = the percent of the population

*Comment on notation:*

1. Since the indicant (i.e., quantified indicator) and population refer to the areas and not to the lines, they are separated by commas making them independent of  $L^{-1}$ , but placed in a parenthesis of which all the contents are operated upon by the  $L^2$  scripts, followed by a colon.

2. Note that in subclassification not all the classes need have subclasses. The number of subclasses for each class is a variable including zero as a value of the variable. Thus, the linear subclass script,  $|_0$ , is 3 for the "roads" subclass of lines, and 0 for the "boundaries" subclass of lines.

3. In a hierarchy of subclasses the successive levels of classes are denoted by successive letters of the alphabet, beginning with the letter of the sector involved.

4. The spatial quantic digits are 21 as these are the exponents in the aggregation of the spatial sector. Since the lines and the areas are aggregated and not multiplied together, their exponents are not added, i.e.,  $(L^{+2} : L^{+1})$  gives  $|^s = t;j;21;p$ , whereas  $(L^{+2}L^{+1})$  gives  $|^s = t;j;3;p$ .

## S. 10



"IN THE EVENING OF HIS LIFE," AS HE PUTS IT, GANDHI HAS VOLUNTARILY WITHDRAWN FROM ACTIVE LEADERSHIP IN THE STRUGGLE FOR INDIAN INDEPENDENCE TO GIVE HIMSELF HEART AND SOUL TO THE CAUSE OF INDIA'S DISINHERITED.

*Ref.:* "Gandhi and Margaret Sanger Debate Birth Control," Portrait by Boris Georgiev, *Asia Magazine*, November, 1936, p. 703.

*Descriptive formula:*  $S_{10} = 'P$

*Quantic number* = 0;0;0;1

*Legend:*

$S_{10}$  = The situation  
records

'P = a particular person (Gandhi)

### S. 11

#### NATIONALITIES AND SECTS:—UNIVERSITY AND LOWER SCHOOLS COMBINED

<i>Nationalities</i>	<i>Students</i>	<i>Nationalities</i>	<i>Students</i>
Abyssinia	2	Trans-Jordan	12
Arabia	1	Turkey	15
Albania	1	U. S. A.	62
Argentina	3	Venezuela	1
* Armenia	6	Zanzibar	2
Assyria	1		
Austria			1421
Brazil	3		
British Empire	19		
Bulgaria	1	<i>Sects</i>	
Czechoslovakia	2	Bahai	17
Denmark		Deist	1
Egypt	78	Druse	69
Equador	2	Hindu	2
France	5	Jew	146
Germany	1	Moslem	451
Greece	19	Zoroastrian	2
India	2	Armenian Catholic	1
Iraq	129	Assyrian	1
Italy	3	Chaldean Catholic	5
Lebanon and Syria	723	Coptic Catholic	1
Lithuania	2	Copt	25
Mexico	6	Gregorian	60
Palestine	238	Greek Catholic	29
Persia	43	Greek Orthodox	261
Poland	6	Maronite	72
Puerto-Rico	1	Nestorian	1
Roumania	1	Protestant	228
Russia	9	Roman Catholics	36
Santo Domingo	3	Syriac Catholics	2
Spain	4	Syriac Orthodox	11
Sudan	14		
Switzerland	1		1421

\* Most of the Armenian students have become Lebanese citizens, and are listed under Lebanon.

*Ref.: American University of Beirut, Leaflet for Visitors, 1933-34.*

Descriptive formula:  $S_{11} = P_{p,q}$ 

Quantic number = 0;0;0;1

Legend:

 $S_{11}$  = The situation  
records $|_{p,q}$  = two classifications (by nation-  
ality and sect) $P$  = the student population of the  
American University of Beirut  
inand  
their classes:  
38 national plurels  
21 sectarian plurels

## S. 12

## SOCIAL DISTANCES BETWEEN RELIGIOUS GROUPS

N = 261 Freshmen of the American University of Beirut

Scale: 0% = maximal friendliness, 50% = indifference,  
100% = maximal hostility

N	Responders	Respondee						
		Greek Orthodox	Jew	Protes- tant	Catho- lic	Moslem Sunnite	Moslem Shiite	Friend- liness
61	Greek Orthodox	(5)	65	25	30	45	45	42
26	Jew	45	(6)	10	32	45	45	39
39	Protestant	35	37	(11)	35	52	60	44
25	Catholic	37	50	25	(2)	35	57	41
82	Moslem Sunnite	37	70	32	37	(2)	30	41
28	Moslem Shiite	40	75	52	42	20	(0)	42
	Popularity	38	62	29	35	41	43	41

"Popularity" =  $^mI_{p':p}$  = weighted average distance towards recipient (respondee)  
group excluding its in-group distance."Friendliness" =  $^mI_{p':p}$  = average distance from expressor (responder) group  
excluding its in-group distance.Ref.: Dodd, Stuart C., "A Social Distance Test in the Near East," *Department of Sociology Year Book* (Typescript), American University of Beirut, Vol. VI, 1933, p. 13.Descriptive formula:  $S_{12} = P_p :: \underline{P}_p : ^mI$ 

Quantic number = 0;1;0;2

Legend:

 $S_{12}$  = The situation

records

: = with a corresponding

 $I$  = indicant of attitude (i.e., social  
distance) in each cell $P_p$  = 6 religious plurels as actors

:: = cross-classified with

 $^m|$  = and means are also stated $\underline{P}_p$  = each other as recipients of ac-  
tion

Comment:

This situation illustrates the *interrelation matrix*, which is believed by the  
author to be the heart of Sociology. When the cells contain a dynamic index  
( $IT^{-1}$ ), it becomes the *interaction matrix*. It is a tool for precise analysis of the

interrelations of persons, or of plurels, which is comparable to the correlation matrix and its summarizing correlation coefficient, which expresses the relationship between the characteristics of a population. Note the quantic formula of interrelation,  $IP^2$ , and of correlation,  $I^2P$ , where the exponent of two denotes the indices, the *relationships* of which, are developed by cross-classification. The systematic development of the forms and uses of the interrelation matrix in Chapter VII is a discovery deduced from the quantic formula, and is believed to be a major contribution of S-theory to Sociology.

In this situation, S. 12, Ch. II, a social distance test in five calibrated steps, ranging from greatest intimacy ("I would be willing to marry —") to greatest hostility ("I wish someone would kill all such persons"), was, for greater reliability, given twice to a student population in Syria. Two-way distances were secured between all pairs of six plurels and one-way distances towards the remaining five plurels. The remarkable finding that emerged, was, that between plurels, friendliness was almost constant, while popularity was highly variable. Note the almost identical means expressing the equal average friendliness of each plurel in the last column, compared with the variable means expressing the dispersion of average popularity of the plurels in the bottom row. This finding stated symbolically and using cross scripts, is:

$$\sigma(MI_{P:P})_P = 1.5 \quad \sigma(MI_{P':P})_P = 10.8 \quad (a)$$

The ratio of in-group to out-group distances (main diagonal cells to all other cells) was 1/10. Its (cross-scripted) formula is:

$$10M(I_{I:I})_P = M(I_{I:I'})_{P:P} \quad (b)$$

*Comment on notation:*

In the "Brief-S" formulae (explained in Chapter IX) of findings, the first one (a) states that in the six plurels expressing attitudes ("expressors"),  $|_P$ , the standard deviation,  $\sigma$ , of the average attitudes,  $M I$ , of one plurel,  $|_{P'}$ , towards the six plurels,  $|_P$ , (including itself) was 1.5, and that among those six plurels the standard deviation,  $\sigma(P)$ , of the mean attitude of the six expressors  $M I_P$ , corresponding to a particular one of the six "receivers,"  $|_{P'}$ , of attitude was 10.8. The dispersion of popularity is seven times the dispersion of friendliness ( $10.8/1.5 = 7.2$ ).

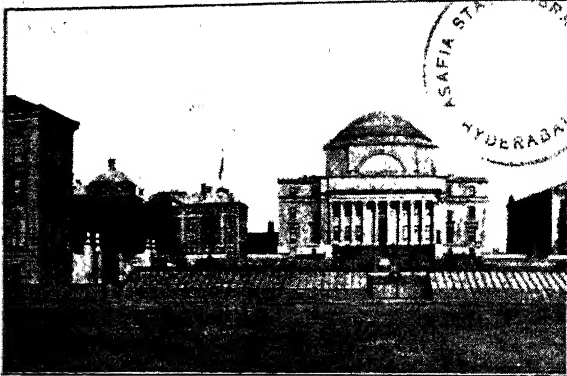
The in-to-out distance ratio formula (b) states that among 6 expressor plurels,  $|_P$ , 10 times the mean,  $10M I$ , of the distance of one of them to itself,  $I_{I:I}$ , equals the mean among the thirty pairs of plurels,  $M|_{P:P}$ , of the distance of any one of them towards another one,  $I_{I:I'}$ .

To one knowing the standardized meaning of the symbols of S-theory, the legend for the situation is sufficient. The findings of trends, or other generalizations revealed in the data, can then be stated, as in (a) and (b) above, without necessity for further legend.

Note that the  $M$  and  $\sigma$  symbols act as operators, directing the reader to compute the mean and standard deviation in the population specified.

## S. 13

## Armament or Education?



© M and L. New York

Columbia University—The Largest in the World  
The photograph shows only a few of its many buildings

Two great universities like this, each a permanent institution, could be established for the cost of one capital ship, which soon becomes obsolete.

The United States has begun the construction of 16 such capital ships.

Ref.: *Facts on Disarmament*, Exhibit, Disarmament Education Committee, Washington, D.C., Card No. 8.

Descriptive formula:  $S_{13} = I_v = w$

Legend:

Quantic number = 0;1;0;0

$S_{13}$  = The situation

records

$|_v$  = universities

= and

$I$  = in an unstated number of dollar units

$|_w$  = warships

an equation between the values

*Comment:*

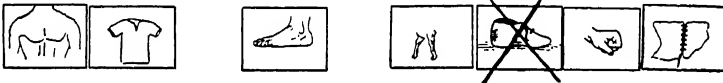
The direct statement equates the two universities with a warship in dollar units, but the question implies other units of "value," however intangible, in the reader's mind, in terms of which he is asked to decide whether the equality should read "greater than" or "less than." His decision, of course, reveals his pacifistic or militaristic set of values. To show explicitly this implied second set of units of "value," a multiple classification script of 2 could be written, covering both members of the equation thus,  $I_1 : (v \geq w)$ , where  $|_1 = 2$ , namely the dollar and the ethical valuations.

It is a psychological phenomena that people are able to make judgments of inequality without consciously conceived units, as in phrases such as, "uglier," "easier," "more spiritual," etc. It is a contribution of S-theory to provide symbols for dealing mathematically with such subjectively quantified qualities, including all kinds of human values. The progressive development of conscious and eventually standardized units for such phenomena may be facilitated by symbols explicitly denoting the unknown ordinal units, and making their current degree of precision and reliability subject to experimental checking.

## S. 14

The next four pages contain 52 lines with 7 pictures in each line. In every line, the first picture is to the second as the third is to one of the four to the right. Cross out the correct one.

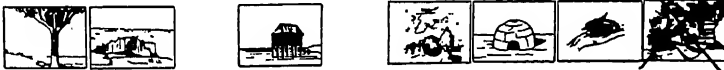
In the line below, "CHEST is to SHIRT as FOOT is to SHOE." The SHOE is therefore crossed out.



In this line, "WING is to BIRD as ARM is to MAN." The MAN is crossed out.



This line is also marked correctly.



On the next four pages, do as many of the 52 items as you can in the time allowed. Work quickly. If you come to one you cannot do, skip it. Cross out the one correct picture of the four at the right in each case.

*Descriptive formula:*  $S_{14} = (I^0_{///} - ///_{(rv, i)})_i$

*Quantic number* = 0;0;0;0

*Legend:*

$S_{14}$  = The situation

with

is a record of

$|_i$  = 3 extra attributes for comparison

$I^0_{///}$  = a ratio, or relation of 2 attributes which are here pictures

$()$  = all repeated in each of

$|_i$  = 3 such "analogy" problems

- = and which is equal to

$///_{rv}$  = a ratio of two other attributes

*Comment:*

The analogy type of intelligence test item requires the completion of a qualitative proportion, whether of pictures or words, by choosing that one element out of several alternatives which makes a perfect proportion, or equality of ratios. This type of item has long been in use by psychologists, but standardized symbols integrating such qualitative thinking with the quantitative techniques of conventional mathematics may prove a major contribution of S-theory.

*Comment on notation:*

1. The equality can be written between indices just as well as between scripts, thus:  $I^0_{///} = I^0_{///(rv, i)}$ . But, without a legend, this does not show as clearly that all the attributes are of one kind.

2. The quantities in the legend can always be written as digits before their respective scripts. The legend then becomes almost unnecessary, except to specify what the qualitative symbols represent in the given situation. Rewriting the formula thus, gives:

$$S = (I^0_{///} - ///_{rv, s_i, s_i})$$

*Legend:*

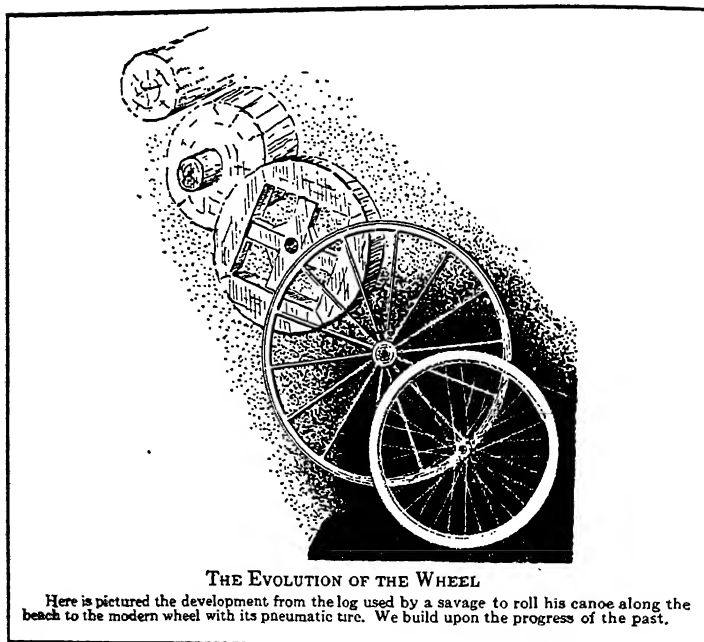
$I^0$  = A picture

The essential pattern of the formula is unaffected by the specific qualitative phenomenon denoted by the  $I^0$ . Instead of "pictures," as here, the analogy might equally well be between objects on a table, words, hymns, movie scenes, historical characters, or any other qualitative characteristics. The legend giving the specific forms of the general concepts is unimportant, compared to the ability which the formula provides to reason precisely about (i.e., to manipulate mathematically) those general concepts and discover relations among them.

## S. 15

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--

## S. 16



*Ref.: Marshall, Leon C., The Story of Human Progress, The Macmillan Company, 1925, p. 164.*

*Descriptive formula:*  $S_{15} = {}^5\text{T}^{-1} : \text{I}^0$

*Quantic number* = 9;0;0;0

*Legend:*

$S_{15}$  = The situation  
records

: = to each of which  
there corresponds

${}^5$  | = 5 dates

$\text{I}^0$  = a style of wheel

$\text{T}^{-1}$  = at unspecified intervals

*Comment on notation:*

1. Since the dates are consecutive, connoting a velocity of evolving, the exponent on the time index is minus one.

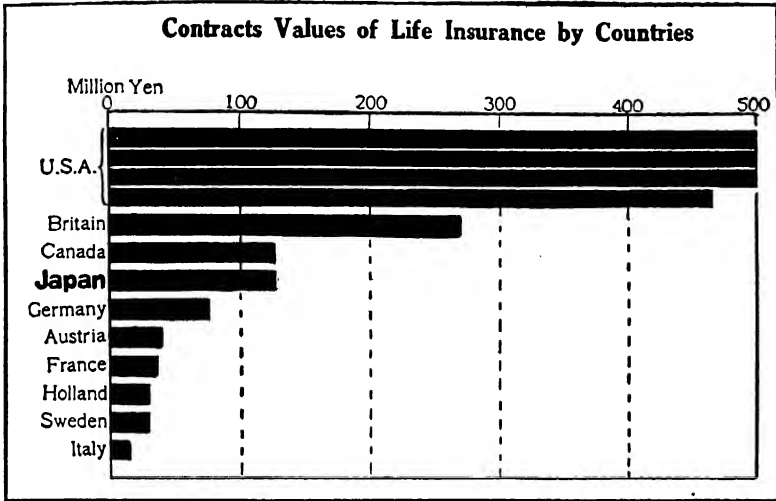
2. As always, the formula can be written with digits inserted, reducing the legend to specifying the particular sort of qualitative characteristics that are denoted by the attribute, thus:

$$S = {}^5\text{T}^{-1} : \text{I}^0$$

*Legend:*

$\text{I}^0$  = A style of wheel.

S. 17



Ref.: Yano, T. and Shirasaki, K., *Nippon, A Chartered Survey of Japan*, Kokusei-Sha, The First Mutual Building, Kyobashi, Tokyo, 1936, Chart 131, p. 335.

*Descriptive formula:*  $S_{17} = \underline{P}_p : I$

*Quantic number* = 0;1;0;1

*Legend:*

$S_{17}$  = The situation  
records

– = (with unspecified population)

: = each with corresponding

$\underline{P}_p$  = 10 national plurels

I = life insurance

*Comment:*

The point of such a distribution as this is to show the absolute amounts of the characteristic possessed by the plurels and the dispersion of it among the plurels. The dispersion is usually measured by the standard deviation, but in the absence of norms from other comparable situations, a sigma of "x million dollars" does not mean much to the reader. It can be made more meaningful by stating the dispersion as a percentage of maximal dispersion. In monopoly the sigma is maximal. Its formula is  $(\frac{\sum_1^p I^2}{p})^{.5}$ . The sum of all the insurances is squared and averaged by dividing by the number of plurels, and the square root taken. At the other extreme is equality among all the plurels with a sigma of zero. The percent of maximal dispersion here is the actual dispersion/maximal dispersion =  $(.569/811) = 70\%$  of the total insurance. Among these 10 nations the situation is 70% of the way from complete equality towards complete monopoly of the total life insurance in force. Since the results of competition vary from equality of the competitors to a monopoly by one of them of all the desideratum competed for, this index of the percentage of monopoly has a direct relation, as will be seen later, to the measurement of competition. (See Eq. 19, and Eqs. 47a, 48, 49, Ch. X for detailed formulae.)

## S. 18

You may travel in a commercial plane with considerable confidence for the records show that fatalities average around 10 million passenger miles per death. That seems like a good many miles, but hour for hour you are in much more danger in a plane than on your front porch or in your own car. Private automobile accidents produce one death about every 15 to 20 million passenger miles. Bus travel rings up about 1 death per 50 million passenger miles unless I miscompute. In fortunate years railways have the excellent record of 1 passenger death in about 1/2 billion passenger miles. The total death due to railroad trains is about 10 times as great as this, that is, 1 person is killed about every 50 million passenger miles, but this includes employees, trespassers and crossing accident victims.

Ref.: Furnas, C. C., *The Next Hundred Years*, Reynal and Hitchcock, 1936, p. 271.

Descriptive formula:  $S_{18} = (\%PL^{-1}T^{-1})_p$

Quantic number = 9;0;9;1

Legend:

$S_{18}$  = The situation

( ) = an index consisting of

records for each of

$\%P$  = the passenger death rate

$_p$  = 5 transportation plurels (airplane, auto, bus, and railroad passengers, and railroad total)

$L^{-1}$  = per mile

$T^{-1}$  = per year

Comment:

The values of the index are:

Airplanes	.000,000,100	= 100 in a billion = $100 \times 10^{-9}$
Autos	.000,000,050 "at best"	= 50 in a billion = $50 \times 10^{-9}$
Busses	.000,000,020	= 20 in a billion = $20 \times 10^{-9}$
R.R. (passengers)	.000,000,002	= 2 in a billion = $2 \times 10^{-9}$
R.R. (all)	.000,000,020	= 20 in a billion = $20 \times 10^{-9}$

Comment on notation:

As noted above, the joint occurrence of a positive and a negative exponent of unity, as in the population rate here, instead of being denoted in the quantic number by the two digits, 1 and 9, is arbitrarily assigned the digit of 1. It is still a primary index with its units merely shifted from an absolute number of persons to a relative number.

## S. 19

## NUCLEATED SOCIAL INSTITUTIONS

Four Type Parts	Family	Church	Government	Business
I. Attitudes and behavior patterns	Affection Love Loyalty Respect	Reverence Loyalty Fear Devotion	Subordination Cooperative-ness Fear Obedience	Workmanship Thrift Cooperation Loyalty
II. Symbolic culture traits, "symbols"	Marriage ring Crest Coat of arms Heirloom	Cross Ikön Shrine Altar	Flag Seal Emblem Anthem	Trade-mark Patent sign Emblem
III. Utilitarian culture traits (real property)	Home Dwelling Furniture	Church edifice Cathedral Temple	Public build-ings Public works	Shop Store Factory Office
IV. Code of oral or written specifications	Marriage li-cense Will Genealogy Mores	Creed Doctrine Bible Hymn	Charter Constitution Treaties Laws Ordinances	Contracts Licenses Franchises Articles of in-corporation

Ref.: Chapin, F. Stuart, *Contemporary American Institutions*, reprinted by permission of Harper and Brothers, 1935, p. 16.

Descriptive formula:  $S_{19} = \underline{P}_p :: I_{i,j}^0$

Quantic number = 0;0;0;1

Legend:

$S_{19}$  = The situation

$I^0$  = "type parts"

is a record of

in

$\underline{P}_p$  = 4 institutional plurels

$|_i$  = 4 classes

- = (with indefinite populations)

each with

:: = cross-classified

$|_j$  = 3 or 4 examples, or subclasses

with their

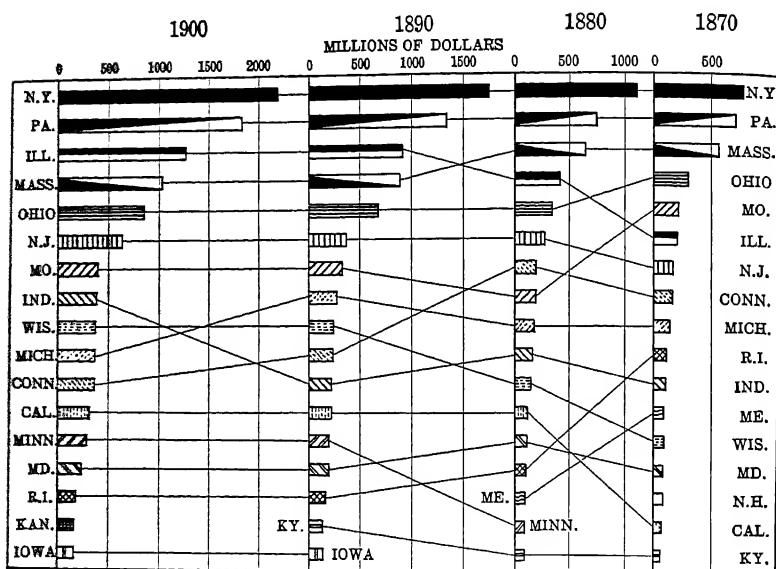
Comment on notation:

1. This classificatory analysis of social institutions is an example of a qualitative matrix. Four kinds of plurels occupy the columns; four kinds of characteristics occupy the rows; and the cells contain listed kinds of subcharacteristics. As usual in qualitative matrices, including contingency tables, both the rows and the columns are commutative, that is, their order or sequence is immaterial. A different arrangement of them would not affect the value of the matrix.

2. Whether to consider the institutions, family, church, government, and business, as plurels (of indefinite size), or as attributes (i.e., qualitative characteristics), is an issue. The former seems more sociologically significant, for a human institution is never existent without people. The  $\underline{P}_p$  is condensed notation, as usual, from  $I^0 \underline{P}$  by Eq. 4, Ch. IV, which denotes the product of the institutional attribute, i.e., its name, and the persons in it.

## S. 20

## VALUES OF PRODUCTS OF MANUFACTURES IN THE SEVENTEEN LEADING STATES, 1870-1900.



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, McGraw-Hill Book Company, Inc., 1914, p. 66.

Descriptive formula:  $S_{20} = {}^{a,2}T^{-1} : \underline{P}_p : (IT^{-1})$

Quantic number = 8;1;0;1

Legend:

$S_{20}$  = The situation

is a record of

$\downarrow$  = 4 intermittent periods

of

$T^{-1}$  = a decade each

${}^{a,2}\downarrow$  = with corresponding terminal dates stated

: = and corresponding to each period

$\underline{P}_p$  = 20 State plurels

: = and for each

I = an indicant of manufacturing production

$T^{-1}$  = per year

Comment:

The fact that three States are eliminated and three others recruited in this list of the seventeen States leading in manufacturing from 1870 to 1900, illustrates the process of *selection* resulting from *competition*.

## S. 21

## THE STANDARD ERROR OF A "SOCIAL FORCE"

## I. DEFINITIONS

In the theory of measurement of social forces certain special cases of frequent occurrence where the population shifts from one date of measurement to the next require the derivation of appropriate standard error formulae.

The theory may be briefly restated<sup>1</sup> in equations as follows: any measurable social change,  $C$ , in a population,  $P$ , may be defined as the difference in mean scores,  $S$ , from surveys or measurements on the dates denoted by subscripts

$$C_{2-1} = S_2 - S_1 = \frac{\Sigma s_2}{P} - \frac{\Sigma s_1}{P} \quad (1)$$

The momentum of a social change may be defined as the product of its time rate in years and the population that is being changed

$$M_{2-1} = PV_{2-1} \quad (2)$$

$$= \frac{PC_{2-1}}{Y_{2-1}} = \frac{P}{Y_{2-1}} (S_2 - S_1) \quad (2a)$$

where  $Y_{2-1}$  is the period from date 1 to date 2 and  $V$  is the velocity, or speed of change, in that period. The acceleration of a social change is definable as the rate of change of the velocity of change

$$A = \frac{V_{4-3} - V_{2-1}}{.5Y_{(4+3-2-1)}} \quad (3)$$

where each velocity, being an average for its period, is taken as representing the mid-date of that period.

The resultant social force which produces a measured change is now definable as that which accelerates the change in a population. It is measurable as the product of the acceleration and the population.<sup>2</sup>

$$F = AP \quad (4)$$

$$= \frac{P}{.5Y_{(4+3-2-1)}} \left( \frac{S_1}{Y_{2-1}} - \frac{S_2}{Y_{2-1}} - \frac{S_3}{Y_{4-3}} + \frac{S_4}{Y_{4-3}} \right) \quad (5)$$

<sup>1</sup> *A Controlled Experiment on Rural Hygiene in Syria*, Dodd, S. C., Publications of the American University of Beirut, Syria, Social Science Series, No. 7, 1934, pp. 336.

Also, *A Theory for the Measurement of Some Social Forces*, Dodd, S. C., Scientific Monthly, Vol. XLIII, No. 1, July, 1936, pp. 58-62.

<sup>2</sup> Force thus defined in terms of its effect is a resultant force, i.e., the residual force after deducting all resisting forces from the total force in the direction of the change observed. This formula defines quantitatively and exactly the "net" force not the "gross" force.

*Descriptive formula:*  $S_{21} = F = P, T^{-1} \text{ ,, } (IT^{-1})$

*Quantic number* = 8;1;0;1

*Legend:*

$S_{21}$ = The situation	$, T^{-1}$ = the time rate of change
$= F$ = defines a societal force	of
as the product of	$\text{,, } T^{-1}$ = a velocity
$P$ = a population	of
and	$I$ = an indicant

*Comment:*

An effective societal force is here defined as a population times its acceleration in some societal change,  $F = PIT^{-2}$ . The concept of the momentum of a societal change is defined "en route" in building up the formula for force, as the product of the population and its velocity of change,  $Mm = P(IT^{-1})$ .

*Comment on notation:*

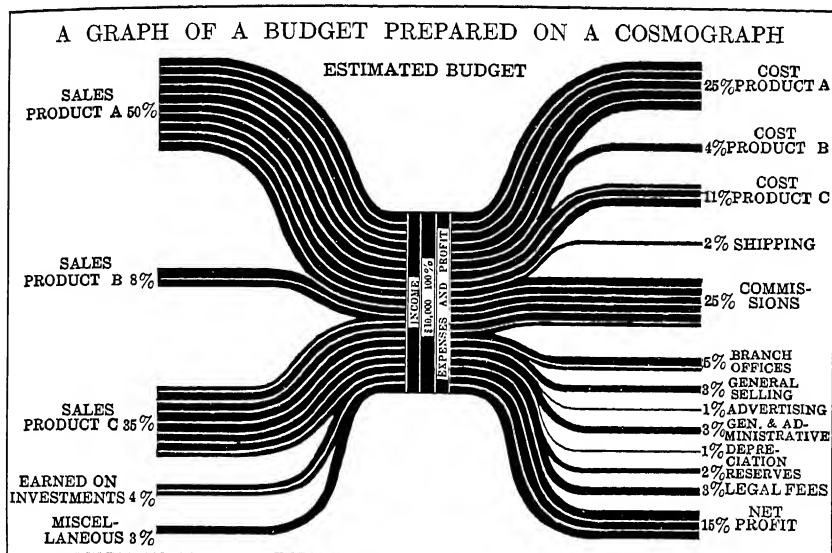
1. Since the data are usually given for dates or periods and not as a smooth curve, mid-dates of periods must be used to state average velocity and acceleration. The date scripts can specify this more precisely for the beginner, yielding the formula as written with less condensation:

$$F = .5 \text{ } ^{'''} + \text{''} - \text{' } T^{-1} P ( \text{ } ^{'''} - \text{''} T^{-1} : I - \text{''} - \text{' } T^{-1} : I ), \text{ where ' , '' , } ^{'''} , \text{ } ^{''} ,$$

denote successive dates bounding the two periods in which the velocity of change must be observed.

This notation states that the amount of change in the indicant between dates ' and '' is divided by the time interval between those dates to get the velocity for the first period. This velocity is subtracted from the velocity of the second period. This change of velocity (in the parenthesis) is divided by the interval between mid-dates. This interval is the average of the last two dates,  $.5 \text{ } ^{'''} + \text{' } T$  (i.e., their mid-date) less the average of the first two dates,  $.5 \text{ } ^{''} + \text{' } T$ . (See Chapter XI for fuller discussion.)

## S. 22



Ref.: Arkin, H., and Colton, R. R., *Graphs: How to Make and Use Them*, reprinted by permission of Harper and Brothers, 1936, p. 113.

Descriptive formula:  $S_{22} = (IT^{-1})_{(I, -//): k}$

Quantic number = 9;1;0;0

Legend:

$S_{22}$  = The situation

$|$  = income

records

and

$T^{-1}$  = an annual

$|$  = expenses plus profit,

$I$  = financial indicant (budget)

each with corresponding

$|$  = in 2 equal classifications

$|_k$  = subclasses

namely:

*Comment on notation:*

1. Since the class-interval script on the indicant is reserved for classes (such as units) that are equal and qualitatively alike, the subclasses here, which are unequal and qualitatively different, are denoted by a subclass script,  $|_k$ .

2. Dynamic indicants which denote events, action, behavior, are denoted by the index,  $(IT^{-1})$ , in distinction from change of a static indicant which exists undiminished no matter how infinitesimal the duration of the observation is made. Change of a static indicant is denoted by  $T^{-1} : I$ , meaning the change of  $I$ -units corresponding to the duration  $T$ . Thus in financial accounting, financial "statements" of assets, liabilities, and proprietorship on two dates are static indicants, whereas expenditure and income, as here, are dynamic indicants. This distinction between dynamic and static indicants has been consistently denoted throughout S-theory and has proved useful. (Compare Eubank, E. E., *Concepts of Sociology*, D. C. Heath and Co., 1931, Part III, subdividing Society Change into Action and Relationship.)

S. 23  
 FACTORIAL MATRIX  $F_4$

$r$ common factors		$n$ specific factors	$n$ error factors
$n$ tests	$a_{11} \ a_{12} \ a_{13} \ \dots \ a_{1r}$	$b_{11}$	$c_{11}$
	$a_{21} \ a_{22} \ a_{23} \ \dots \ a_{2r}$	$b_{22}$	$c_{22}$
	$a_{31} \ a_{32} \ a_{33} \ \dots \ a_{3r}$	$b_{33}$	$c_{33}$
	$\dots \ \dots \ \dots \ a_{jm} \ \dots$	$b_{jj}$	$c_{jj}$
	$a_{n1} \ a_{n2} \ a_{n3} \ \dots \ a_{nr}$	$b_{nn}$	$c_{nn}$

Ref.: Thurstone, L. L., *The Vectors of Mind*, Univ. of Chicago Press, 1935, Table 2, p. 57.

*Descriptive formula:*  $S_{23} = I_i; j; k$

*Quantic number* = 0;1;0;0

*Legend:*

$S_{23}$  = The situation  
                     records

$|_j$  = 3 types of "factors" —com-  
                     mon, specific, and error

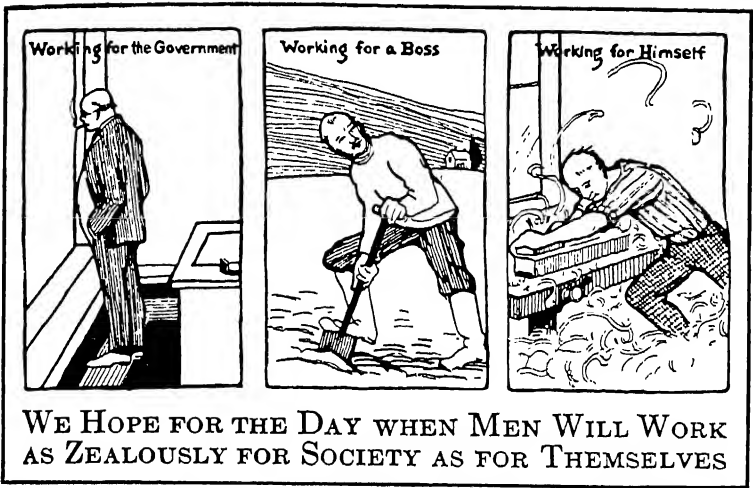
$I_i$  =  $n$  tests  
                     each subclassified into

                    subdivided into  
 $|_k$  = the  $2n + r$  factors

*Comment:*

This S illustrates the components (which are also called factors) derived by psychologists from the intercorrelation coefficients of a set of tests given to one population. These types of analyses which involve complicated matrix and determinantal algebra to derive their centroids, principal axes, or other type of components, can readily be handled in the notation of S-theory. Thus, the fact that the oblong matrix, here presented, is subdivided into three submatrices is very simply specified by the colon symbol of subclassification.

S. 24



Ref.: Marshall, Leon C., *The Story of Human Progress*, The Macmillan Company, 1925, p. 411.

Descriptive formula:  $S_{24} = {}^pP : P_p' : {}^i(IT^{-1})$       Quantic number = 9;1;0;2

Legend:

$S_{24}$ = The situation	showing
records	
${}^pP$ = 3 particular but unnamed	${}^i()$ = an ordinal index
workers (i.e., types)	of
and each in relation to	$IT^{-1}$ = zeal or production per period
$P_p'$ = a type of employing person or	in unspecified units
plurel	

Comment:

This situation is a partial interrelation matrix of the "asymmetric principal pairs" type. Full interrelation (see S. 12) involves the relation of every one of P persons (or plurels) in rows, cross-classified against every one of those P persons in columns. Asymmetric interrelations are from the submatrices of quadrants 1 and 3 at right, where (as here) employees are shown in relation to a different set of persons, i.e., employers. But in this situation, S. 24, only the pairs in the principal diagonal of the  $3 \times 3$  submatrix, quadrant No. 1, are involved, hence it is specifiable as the asymmetric principal pairs part of the interrelation matrix. All interrelations of people in respect to one characteristic at a time can be systematically represented by the interrelation matrix, including parts of it, and by summarizing indices of parts or of the whole.

	'P	"P			*P
'P					
"P		#2		#1	
		#3		#4	
*P					

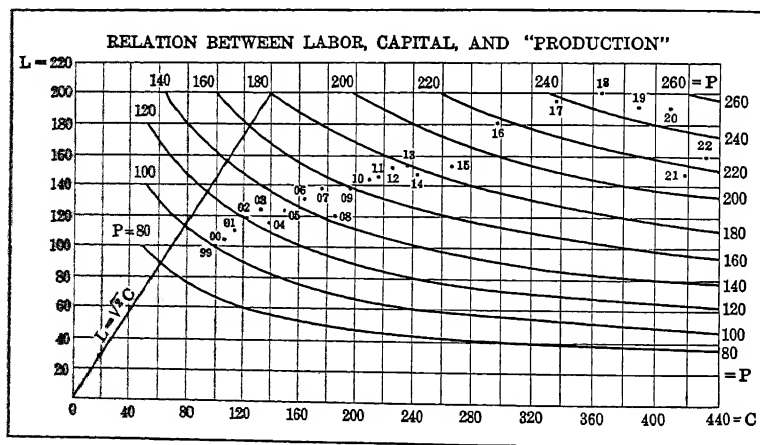
*Comment on notation:*

1. Note that the underlining symbol of indefiniteness when coupled with the person script or any singular script can denote a type (a particular but unidentified case, therefore any one of its kind). This definition of a type, as substitutable for every one of the members of the class typified, is the most inclusive mathematical definition of a mean,  $M$ , of  $n$  variates,  $x$ .

$f(x_1, x_2, \dots, x_n) = F(M, M, \dots, M)$ . (See Dodd, E. L., University of Texas, Bulletin 3606, 1936, p. 32.)

2. The "zeal" of work, expressed in pictures, is in ordinal numbers, i.e., numbers in a series each greater than all on one side of itself, and less than all on the other side, but with no assertion that the successive numbers are equal steps. Ranks are such ordinal numbers contrasted with cardinal numbers which are multiples of equal interchangeable units. Ordinal indicants may be denoted, as here, by the case script specifying points on a scale, without the connotation of equal intervals contained in the class-interval script.

## S. 25



Ref.: Cobb, Charles W., "Contour Lines in Economics," *Journal of Political Economy*, Univ. of Chicago Press, Vol. XXXVI, Feb.-Dec. 1929, Chart I, p. 226.

Descriptive formula:  $S_{25} = {}_P P :: {}_I I : {}_I I$

Legend:

$S_{25}$  = The situation  
records

${}_P P$  = labor population  
:: = cross-classified with

Quantic number = 0;2;0;1

${}_I I$  = an indicant of capital  
: = with a corresponding  
 ${}_I I$  = indicant of production plotted

*Comment:*

The situation as presented in the graph is analyzed into the *descriptive* S-formula above. It is a correlation expressing the dependence of one variable (production) on two other variables jointly. Study of the text of the article, however, reveals that the graph's curves are not observational data, but the plotting of the calculative formula:  ${}_1I = {}_1I^3P^3$ , which expresses a hypothesis as to the functional relation of two of the factors of economic production. In this equation the correlation is perfect,  $r = 1.0$ , i.e., production is exactly equal to the function on the right without allowance for any error of estimate.

*Comment on notation:*

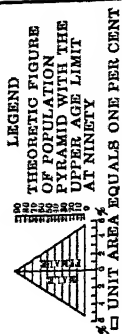
1. The  ${}_1I$ ,  ${}_1I$ , and  ${}_pP$  of S-notation represent the P, C, L respectively of the notation on the graph.

# CHANGES IN POPULATION BY AGE AND SEX

IN OAK PARK AND SELECTED CENSUS TRACTS  
OF CHICAGO FOR 1910, 1920, 1930.

## NOTE

OBSERVED DISTRIBUTION OF  
EACH AREA COMPARED WITH  
THEORETICAL DISTRIBUTION  
OF POPULATION, EACH OF WHICH  
EQUALS ONE HUNDRED PER CENT



## 1930 SERIES



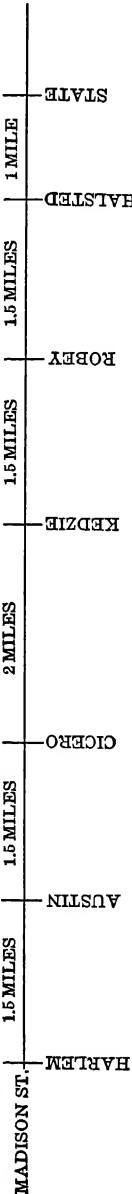
## 1920 SERIES



## 1910 SERIES



1910  
DATA NOT  
AVAILABLE



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill Book Company, 1933, p. 296.

S. 26 (*Continued*)

*Descriptive formula:*  $S_{26} = {}^tT^{-1} : ({}^zT^1 : P_p : q)_r$       *Quantic number* = 91;0;0;1  
*Legend:*

$S_{26}$ = The situation	${}^tT^1$ = 18 5-year ages,
is a record	${}^z $ = up to 90 years,
${}^tT^{-1}$ = at each of 3 decennial census dates	for each of
and for each of	$ _p$ = the 2 sexes
$ _r$ = 6 Chicago census tracts (plurels)	with corresponding
of a population pyramid or matrix	$P$ = frequencies of persons
of	in
	$ _q$ = 2 classifications—actual and theoretic

*Comment:*

The hypothesis that population pyramids are normally isosceles is checked in this set of urban pyramids revealing them to be non-normal, as they come from areas with little genetic population and mostly from an occupational (commuting) population with its excess of males and of adults of gainfully-occupied ages.

The tendency towards isosceles pyramids is measurable by the formula  $T \cdot P < 0$ , which states a negative ( $<0$ ) correlation ( $\bullet$ ) between age and frequency of people. When the pyramid is exactly isosceles the correlation coefficient is perfect, i.e.,  $T \cdot P = -1.00$ .

*Comment on notation:*

1. This is a fifth-degree matrix. The 1512 population cell entries are in 2 reality arrays,  $|_q$ , by 2 sex arrays,  $|_p$ , by 18 age arrays,  $|_t$ , by 7 tract arrays,  $|_r$ , by 3 date arrays,  $|_t$ . ( $2 \times 2 \times 18 \times 7 \times 3 = 1512$ .)

2. The difference between the positive and negative exponents of unity on the temporal indices distinguishes change from duration (which is also change but a relatively slow one). A common initial date connotes a velocity of change denoted by  $T^{-1}$ , whereas a common terminal date such as the present moment, and variable initial dates such as the birthdays of a population, connote durations up to the present,  $T^{+1}$ . A single population remeasured at different ages, as in growth studies, connotes velocity of change and is denoted by  $T^{-1}$ .

## S. 27

## A TENSION THEORY OF SOCIETAL ACTION

The ratio of the average intensity of desire (D) of a population (P) for a specified desideratum to the amount of the desideratum available (V) equals and defines a "coefficient of societal tension" (E)

$$E = \frac{PD}{V} \quad (\text{the simple tension equation})$$

A matrix of such equations repeated for varying dates or periods ( $\dagger 1$ ) for various plurels ( $\dagger p$ ) with either persons (P) or sub-plurels ( $\dagger p$ ) as units, for various values ( $\dagger v$ ) with varying exponents ( $\dagger e$ ):

$${}_{\dagger p}({}_{\dagger v}PD = {}_{\dagger e}VE)_{\dagger p, \dagger v} \quad (\text{the matrix tension equation})$$

can describe quantitatively in symbols many historical as well as theoretical, actual as well as planned, societal situations and developments. Derived equations serve to define many societal processes and their corresponding relations such as "conflict," "competition," "accommodation," "co-operation," "association," "differentiation," "progress," and other sociological concepts.

*Descriptive formula:*  $S_{\dagger p} = {}_{\dagger e}T^{-1} : {}_{\dagger p}(PD = VE)_{\dagger p, \dagger v}$       *Quantic number* = 9;1;0;1

*Legend:*

$S_{\dagger p}$ = The situation	the indices being a product of
records	
(=) = an equality of indices	P = a population
on	times
$\dagger T$ = a series of dates and/or	D = its average intensity of desire
periods	equaling the product of
among	V = the quantity of the desider-
$\dagger p$ = aggregations of plurels	ata
$::$ = cross-classified with	
$\dagger v$ = a set of indicants of desider-	times
ata	E = an equilibrating coefficient
$\dagger e$ = with varying pattern of ex-	
ponents,	

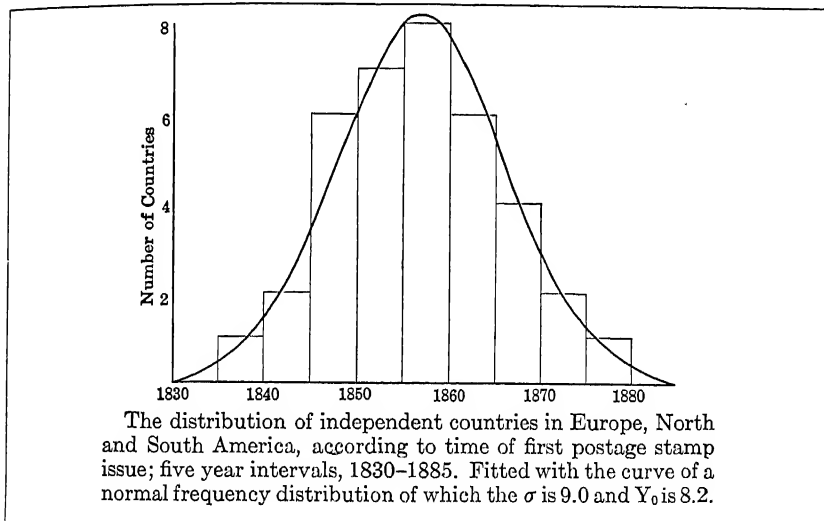
*Comment:*

This tension, or equilibrium, theory of societal action is an example of the fruitfulness of objective symbols in theoretical research. It was the first step in the development of the S-theory. By manipulation of this tension theory equation, more than thirty of the societal processes which involve desiderata (i.e., human desires of any kind) have been defined. Thus "competing," illustrated in S. 20, is an example where the standard deviation of the gains and losses among a set of competing plurels in respect to a desideratum measures the competing as a percentage of the maximal possible competing. Similarly, "effective co-operating" may be defined as the increase of the desideratum, V, resulting from the efforts of the persons or plurels desiring it. "Effective conflict" is definable as the decrease of the population, P, resulting from competitors

ceasing to strive exclusively for the limited desideratum and turning to eliminating some of the population competing for that desideratum.

Generalizing these equations of this tension theory led to developing the basic S-theory equation (Eq. 50, Ch. II) and consequently all the equations in this volume.

## S. 28



Ref.: Pemberton, H. Earl, "The Curve of Culture Diffusion Rate," *American Sociological Review*, Vol. I, No. 4, Aug. 1936, p. 552.

Descriptive formula:  $S_{28} = {}_tT^{-1} : (I^0_p PT^{-1})$

Quantic number = 8;0;0;1

Legend:

$S_{28}$  = The situation  
records

${}_pP$  = a frequency of national plurels  
adopting

$|$  = 11 periods

$T^{-1}$  = of 5 years each

$I^0$  = postage stamps

$'$  = from 1830 on

$T^{-1}$  = per period

$:$  = corresponding to which are

Comment:

These data on the rate of diffusion of a culture trait are fitted by a normal probability curve. The hypothesis was that since the adoption of postage stamps was determined by a large number of small and more or less independent causes cumulatively operating to bring governments to such a decision, the normal curve, which is the resultant of many small, independent elements, would describe the data well. The goodness of the fit of the curve to the data is the degree of verification of the hypothesis, that this cultural diffusion follows the normal probability law.

The equation of the fitted normal curve above in the S-notation of the legend is:

$$p| = \frac{8.2}{2.718^{T^2/162}}$$

## S. 29

## ANNUITY RATES

Age Last Birth- day	Cost of an Annuity of \$100 per Annum Payable Yearly		Amount of Annuity, Payable as Indicated, Which \$1000 Will Purchase					
			Males			Females		
	Males	Females	Yearly Pay- ment	Half- Yearly Pay- ment	Quar- terly Pay- ment	Yearly Pay- ment	Half- Yearly Pay- ment	Quar- terly Pay- ment
10	\$2009	\$2035	\$49.78	\$24.58	\$12.21	\$49.14	\$24.27	\$12.05
11	1998	2026	50.05	24.72	12.28	49.36	24.38	12.11
12	1987	2015	50.33	24.85	12.35	49.63	24.51	12.18
13	1975	2005	50.63	25.00	12.42	49.88	24.63	12.24
14	1963	1994	50.94	25.15	12.49	50.15	24.76	12.30
15	1951	1983	51.26	25.30	12.57	50.43	24.90	12.37
16	1939	1972	51.57	25.46	12.65	50.71	25.04	12.44
17	1927	1962	51.89	25.61	12.72	50.97	25.16	12.50
18	1915	1951	52.22	25.77	12.80	51.26	25.30	12.57
19	1903	1940	52.55	25.93	12.88	51.55	25.45	12.64
20	1891	1930	52.88	26.10	12.96	51.81	25.58	12.70
21	1879	1919	53.22	26.26	13.04	52.11	25.72	12.77
22	1866	1908	53.59	26.44	13.13	52.41	25.87	12.85
23	1853	1897	53.97	26.62	13.22	52.71	26.01	12.92
24	1840	1885	54.35	26.81	13.31	53.05	26.18	13.00
25	1827	1874	54.73	27.00	13.40	53.36	26.33	13.08
26	1813	1861	55.16	27.20	13.51	53.73	26.51	13.16
27	1798	1849	55.62	27.43	13.62	54.08	26.68	13.25
28	1784	1836	56.05	27.64	13.72	54.47	26.87	13.34
29	1769	1823	56.53	27.87	13.84	54.85	27.06	13.43
30	1753	1810	57.05	28.12	13.96	55.25	27.25	13.53
31	1737	1796	57.57	28.38	14.08	55.68	27.46	13.63
32	1720	1782	58.14	28.65	14.22	56.12	27.67	13.74
33	1703	1767	58.72	28.94	14.36	56.59	27.90	13.85
34	1685	1752	59.35	29.24	14.51	57.08	28.14	13.97
35	1667	1736	59.99	29.55	14.66	57.60	28.39	14.09
36	1648	1720	60.68	29.89	14.83	58.14	28.65	14.22
37	1629	1704	61.39	30.23	15.00	58.69	28.92	14.35
38	1609	1687	62.15	30.60	15.18	59.28	29.21	14.49
39	1588	1669	62.97	31.00	15.38	59.92	29.52	14.65
40	1567	1652	63.82	31.41	15.58	60.53	29.82	14.79
41	1545	1634	64.72	31.85	15.79	61.20	30.14	14.95
42	1523	1616	65.66	32.30	16.02	61.88	30.47	15.11
43	1501	1597	66.62	32.77	16.24	62.62	30.83	15.29
44	1478	1578	67.66	33.27	16.49	63.37	31.19	15.47
45	1455	1559	68.73	33.78	16.74	64.14	31.57	15.65

## S. 29 (Continued)

## ANNUITY RATES

Age Last Birth- day	Cost of an Annuity of \$100 per Annum Payable Yearly		Amount of Annuity, Payable as Indicated, Which \$1000 Will Purchase					
			Males			Females		
	Males	Females	Yearly Pay- ment	Half- Yearly Pay- ment	Quar- terly Pay- ment	Yearly Pay- ment	Half- Yearly Pay- ment	Quar- terly Pay- ment
46	1431	1540	69.88	34.34	17.02	64.94	31.95	15.86
47	1407	1519	71.07	34.92	17.30	65.83	32.38	16.07
48	1382	1499	72.36	35.54	17.61	66.71	32.81	16.24
49	1357	1478	73.69	36.18	17.92	67.66	33.27	16.49
50	1331	1456	75.13	36.87	18.26	68.68	33.76	16.73

Ref.: *Prospectus*, Sun Life Assurance Company of Canada, p. 25.

Descriptive formula:  $S_{29} = {}^tT^{+1} : \underline{P}_p : {}_uT^{-1} : \%I$       Quantic number = 19;1;0;1

Legend:

$S_{29}$  = The situation  
records

${}^tT$  = 41 ages

'| = beginning at 10

$\underline{P}_p$  = for each sex

for each of which there cor-  
responds

${}_uT^{-1}$  = 3 periods of premium pay-  
ment

with corresponding

$I$  = premiums

$\%|$  = in a relative number of dol-  
lars

Comment:

The table represents the application of well-verified laws of probability. The field of insurance is, perhaps, one of the best refutations of the pessimistic charge that the social sciences will never be able to develop quantitative laws usable for practical prediction of human affairs.

Comment on notation:

1. The two classifications of time, the indices  ${}^tT$  and  ${}_uT$ , illustrate again the convention in S-theory of considering ages on a given date as durations prior to that date and denoting this with a positive exponent, while the periods of premium payments are considered as specifying the velocity,  $(T^{-1})$ , in three class-intervals, of payment of the total premium.

2. When an indicant is expressed as a percent of a constant base indicant, as in index numbers, the symbol,  $\%|$ , denotes this by directing the operator to divide the variable values of the indicant  $I$  by a constant or single particular value of itself.

## S. 30

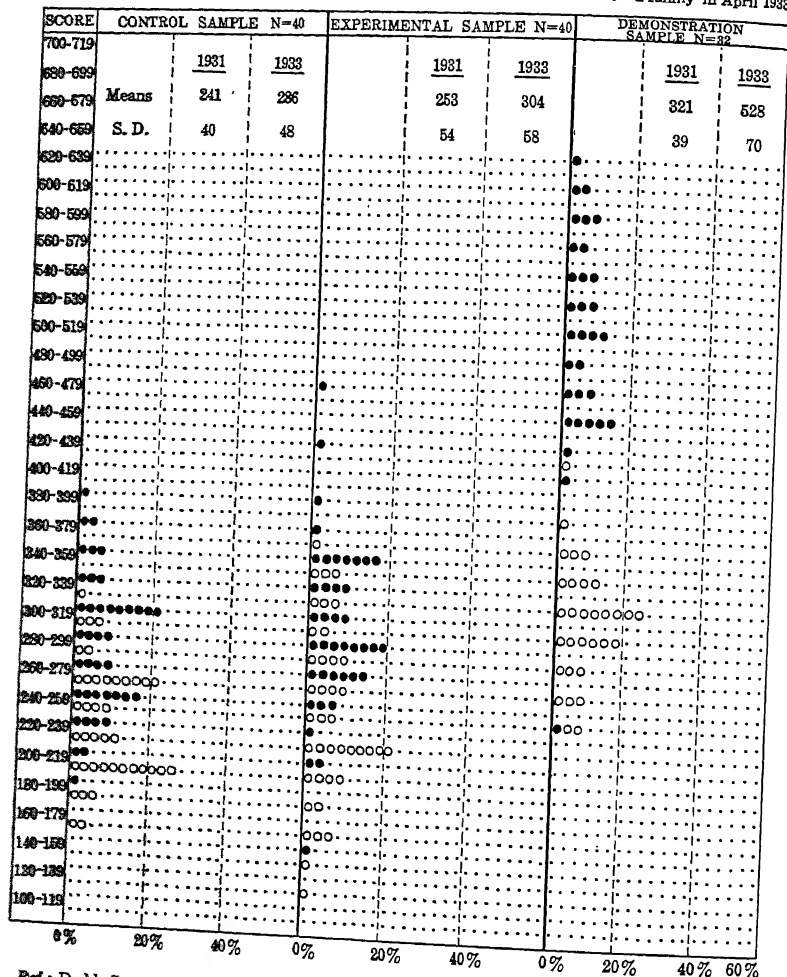
## SCALE SCORES

COMPARATIVE DISTRIBUTION OF SECTION SCORES  
AT THE BEGINNING AND END OF THE EXPERIMENT IN  
THE EXPERIMENTAL, CONTROL & DEMONSTRATION SAMPLES

Maximum Score = 1000

O = a family in April 1931

● = a family in April 1933



Ref.: Dodd, Stuart C., *A Controlled Experiment on Rural Hygiene in Syria*, American Press, Beirut, 1934, p. 188.

*Descriptive formula:*  $S_{30} = ({}^tT^{-1} : \sigma_{m,z}, {}_iI : {}_pP)_p$

*Quantic number* = 9;1;0;1

*Legend:*

$S_{30}$  = The situation  
records

${}_iI$  = a hygiene indicant (score)

$|_p$  = 3 experimental plurels

among

( $|_p = |, |'', |'''$ )

$|_p$  = the families

in each of which on

and

${}^tT$  = 2 specified dates

$\sigma_{m,z}$  = the sigma, mean, and maximum scores are also stated

there is stated the distribution  
of

*Comment:*

A clean-cut example of the experimental method applied upon whole communities is statistically summarized in this S-situation dealing with hygienic progress. By using cross scripts, the findings can be compactly stated without further legend. Thus the mean progress in each plurel is:  ${}_iI$  = 20%,  ${}_iI''$  = 18%,  ${}_iI'''$  = 65%, and the reliability of each gain as a ratio to its standard error is:

$$\sigma({}_iI) = 4, \quad \sigma({}_iI'') = 8, \quad \sigma({}_iI''') = 16$$

showing excellent reliability since these significant ratios all exceed the conventional value of 3.

S. 31

The planned estimates for freight and passenger traffic and capital investments were greatly exceeded by 1932, though in many other respects the attainments lagged considerably. The following table shows the achievements for the period in respect of goods and passengers carried and the capital investments made in relation to the estimates laid down in the Plan:

Year	Freight Million Tons		Passengers Millions		Capital Investment Million Roubles	
	Plan	Fulfillment	Plan	Fulfillment	Plan	Fulfillment
1929	165.0	187.6	302.3	365.2	518	873
1930	185.0	238.7	337.5	557.7	743	1242
1931	210.0	258.3	380.8	723.7	950	1960
1932	240.3	267.9	416.7	967.1	1187	3000

Ref.: U.S.S.R. Handbook, Victor Gollancz Ltd., London, 1936, p. 138.

*Descriptive formula:*  $S_{31} = ({}^tT^{-1} : ({}_iI, P), {}_jT^{-1})$

*Quantic number* = 8;1;0;1

*Legend:*

$S_{31}$  = The situation  
records

and

${}^tT$  = 4 annual periods

$P$  = the number of passengers  
 $T^{-1}$  = per year

$'$  = beginning in 1929

$|_j$  = in 2 classifications, planned  
and fulfilled

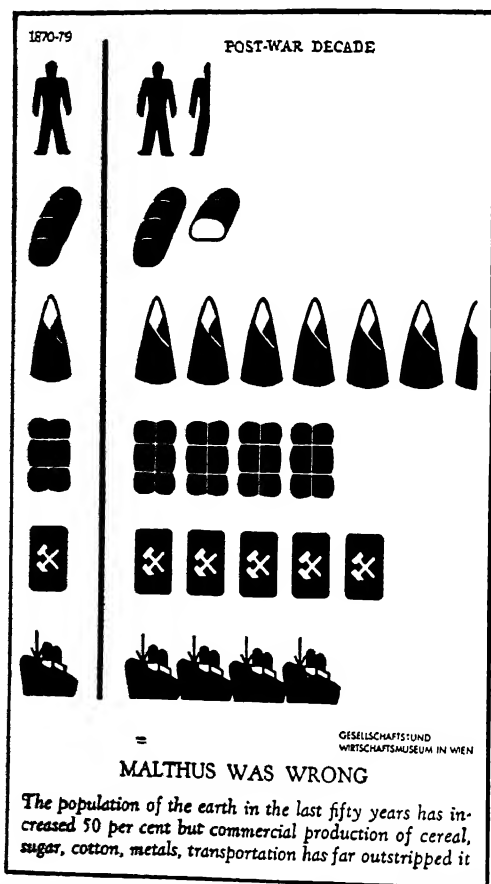
to each of which there corresponds

${}_iI$  = an indicant of freight and one  
of capital

*Comment:*

This extract from the Russian Five-Year Plan expresses, in a matrix, the degree of societal control actually achieved as compared with that planned. It illustrates the increasing precision in predicting major societal phenomena as well as the increasing precision in subsequently verifying such predictions.

S. 32



Ref.: Neurath, Otto, "World Planning and the U. S. A." *Survey*, Vol. LXVII, No. 11, March 1, 1932, p. 629.

*Descriptive formula:*  $S_{22} = T^{-1} : P < (iIT^{-1})_i$

*Quantic number* = 8;1;0;1

*Legend:*

$S_{22}$  = The situation  
is a record of

$T^{-1}$  = a sixty-year period,

$i$  = beginning in 1870  
and a corresponding

$P$  = increase of population

$<$  = which is less than  
the increase in each of

$i_i$  = 5 indicants of production

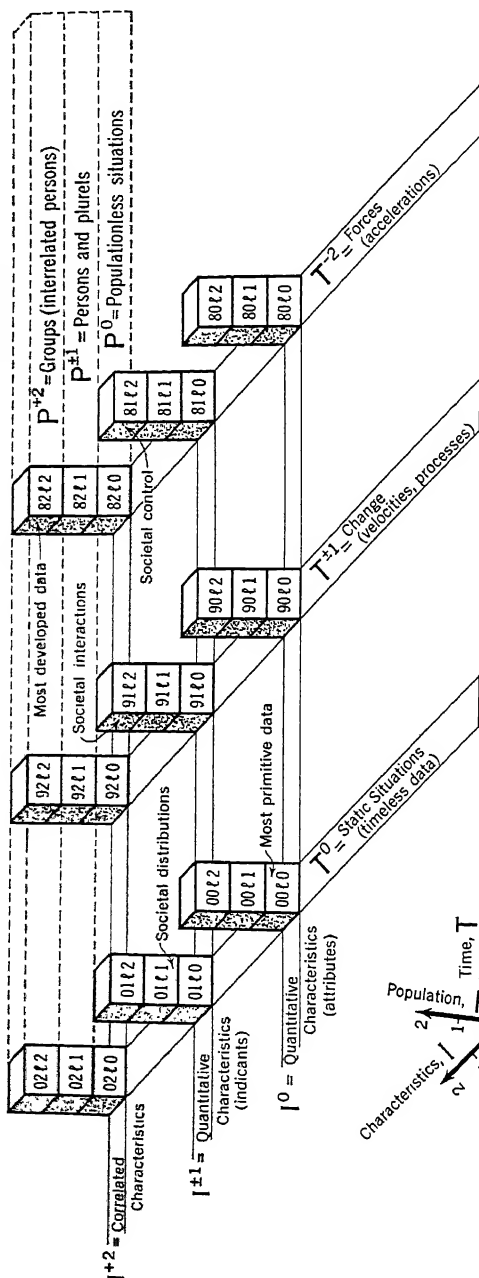
$T^{-1}$  = per decade

in

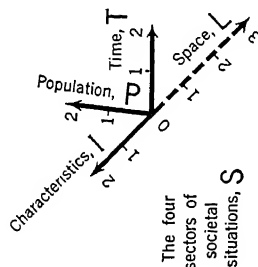
,| = class-intervals equal to the  
1870-79 production*Comment:*

The Malthusian theory requires that the increase of population should tend to exceed (or may equal, if the "checks" are effective) the increase in the food supply. The reversal of the inequality sign above, in stating that the increased production exceeded the population increment, disproves the Malthusian theory in a world situation where scientific technology and birth control have developed.

# THE "QUANTIC SOLID" OF S - THEORY a Diagram of the Quantic Classification of Societal Phenomena



Each block represents one class of quantitatively expressed societal situations. The "quantic number" identifying each block is composed of the exponents on the four indices in the quantic formula:  $S = T^1 P^1 L^1 P^0$



S. 33 (*Continued*)

*Descriptive formula:*  $S_{33} = \bar{T}^t :: \bar{I}^i :: \bar{L}^l :: \bar{P}^p$

*Quantic number* =  $t; i; l; p$

*Legend:*

$S_{33}$  = The situation  
records

- = where base indices are of in-  
definite amounts

T = the temporal

all composing

$|^t$  = exponents 0, 1, -1, and -2

- = the quantic solid, a geometric  
representation of the quantic  
classification

:: = cross-classified with

$I^i$  = the indicatory exponents 0, 1,  
and 2

:: = cross-classified with

$P^p$  = the populational exponents 0,  
1, and 2

*Comment:*

The categories of the fundamental quantic formula, which classifies all quantitative societal phenomena, are here geometrically represented as cubes. The phenomena symbolized by successive arrays of these cubes are systematically studied in the successive chapters of this volume. Each wooden cube bears a quantic formula and its quantic number. The fourth variable, space,  $L^l$ , is necessarily omitted in this three-dimensional diagram.

This geometric representation of the exponents and that of S. 35, diagramming indices and descripts, become consistent and mutually complementary if the linear scales in S. 35 are drawn on logarithmic paper. Then successive equal increments as of one, two, or three inches on the paper along the vectors of S. 35 represent the successive exponents  $|^1, |^2, |^3$ , etc., which are the successive arrays of blocks in S. 33 above.

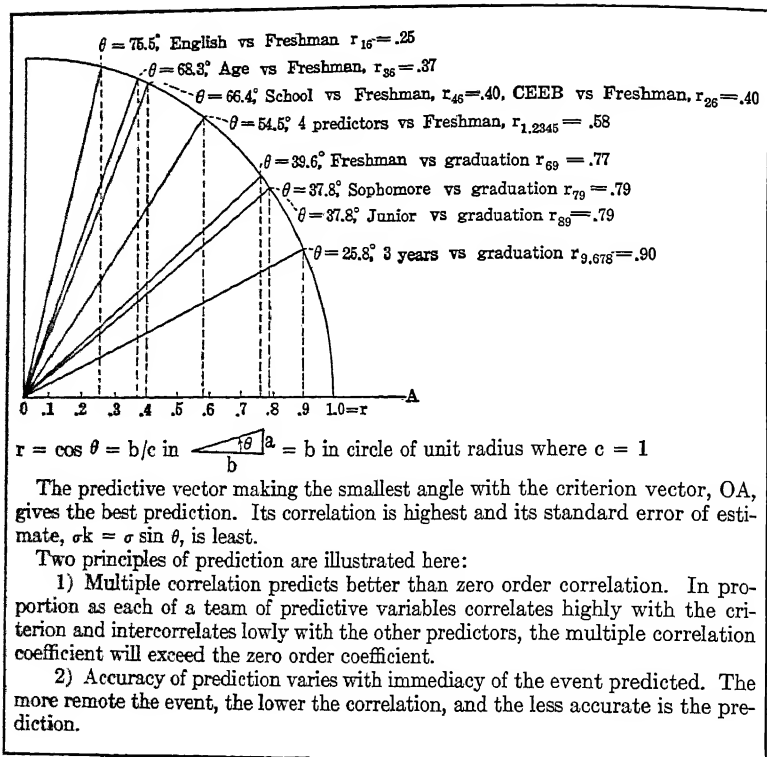
## S. 34

## PREDICTIONS FROM CORRELATIONS GRAPHED AS COSINES

In a large private University in the eastern States the following correlations were observed. The variables were:

1. English = College Entrance Examination Board (C.E.E.B.) English examination
2. C.E.E.B. = College Entrance Examination Board, weighted average
3. Age = candidates age at admission to Freshman year
4. School = quartile rank in secondary school class
5. Aptitude = Scholastic Aptitude Test of the C.E.E.B. (Verbal section)
6. Freshman = rank in 7 class-intervals for all Freshman courses
7. Sophomore = rank in 7 class-intervals for all Sophomore courses
8. Junior = rank in 7 class-intervals for all Junior courses
9. Graduation = rank in 6 class-intervals upon graduation

## S. 34 (Continued)



Ref.: Stuart C. Dodd's Files, Notes from Report to Harvard College Deans, 1935.

Descriptive formula:  $S_{24} = (\sigma I \cdot \sigma I)_{i,j}$

Quantic number = 0;2;0;9

Legend:

$S_{24}$  = The situation

is a record of

— = the geometric interpretation

of

(•) = the correlation coefficients

of

$I_i$  = 5 indicants of college achievements

with their corresponding

$I_j$  = 9 indicants of prediction

$\sigma$  = (when in sigma units)

Comment:

The correlation coefficients are here graphed as cosines of the angles between the unit-vectors representing the indicants correlated. Each coefficient is the scalar product of two vectors,  $\bar{I}_i \cdot \bar{I}_j$ , whose scalar lengths have been first expressed in sigma units and later averaged by dividing by the scalar, P.

The purpose of the research graphed here is to determine the predictive vector which makes the smallest angles with the horizontal criterion vector, since the smaller the angle the greater the correlation and the more efficient the prediction.

## S. 35

VECTORIAL INTERPRETATION OF S-THEORYPopulation SectorShowing 2 vectors,  $\bar{P}_p$ 

$P = \text{a plurel}$   
 $(P = \text{a person})$

Indicator Sector

Showing 3 coplanar vectors,  $\bar{I}_i$   
 $OA \perp OF$  i.e.  $\bar{I}_i$  is normal to  $\bar{I}_{iii}$

$I = \text{a Class-interval}$   
 $i = \text{a Point}$

Length SectorShowing 1 vector cubed,  $\bar{L}^3$ Time SectorShowing 2 collinear vectors  $\bar{T}_t$ 

Geometrically: 1) A vector represents an index,  $((I') = T, I, L, P)$

2) The exponent represents the number of mutual perpendiculars  
 (Note  $L^3 = L' \times L'' \times L'''$ )

3) The class script represents the number of vectors ( $|_s = |_t, |_i, |_l, |_p$ )

4) The class-interval script represents the number of line sects  
 ( $|_s = |_t, |_i, |_l, |_p$ )

5) The case script represents the number of specified points  
 ( $|_s = |_t, |_i, |_l, |_p$ )

Ref.: Department of Sociology Year Book (Typescript), American University of Beirut, Vol. VIII, p. 279.

Descriptive formula:  $S_{35} = \bar{t}T_t^{+1}, \bar{i}I_i^{+1}, \bar{l}L_l^{+3}, \bar{p}P_p^{+1}$       Quantic number = 1;1;3;1

Legend:

$S_{35}$  = The situation  
 records

$L_l^3$  = 1 length index, cubed  
 and

— = the vectors  
 representing





$P_p$  = 2 population indices,  
 and

$T_t$  = 2 time indices,  
 $I_i$  = 3 indicant indices,

$|_s$  = their points and sects

Comment:

This diagram <sup>39</sup> presents the geometrical meaning of indices, as vectors, and of their descripts. In addition to the five interpretative comments below the diagram, the following equivalences in vectorial, geometric, and statistical terms may be noted:

<i>Some equivalences</i>	<i>in</i>	<i>vector algebra</i> <sup>40</sup>
(6) Scalar product of vectors:		$\bar{I}_I \cdot \bar{I}_J = I_I I_J \cos \theta$
(7) Scalar product of unit vectors:		$\bar{I}_I \cdot \bar{I}_J = \cos \theta$
(8) Scalar product of perpendicular vectors, $\bar{I}_I \perp \bar{I}_J$ :		$\bar{I}_I \cdot \bar{I}_J = 0$
(9) Scalar product of collinear vectors, $\bar{I}_I // \bar{I}_J$ :		$\bar{I}_I \cdot \bar{I}_J = I_I I_J$
(10) Scalar product of collinear and equal vectors, $\bar{I}_I = \bar{I}_J$ :		$\bar{I}_I \cdot \bar{I}_J = I_I^2$
(11) Sum of vectors:		$\bar{I}_I + \bar{I}_J = \overline{OC}$
(12) Difference of vectors:		$\bar{I}_I - \bar{I}_J = \overline{BA}$
<i>geometry</i>	<i>and</i>	<i>statistics</i>
(6) = area  OBEF <sup>41</sup>		= $\sigma_I \sigma_J r_{IJ}$ , the covariance
(7) = area  OBEF		= $r_{IJ}$ (since $\sigma_I = \sigma_J = 1$ )
also = length OD (since $OA = OM = 1$ )		
(8) = zero area (since $\cos 90^\circ = 0$ )		= 0 (since $r = 0$ )
(9) = area  OBEF (since $\cos 0^\circ = 1$ )		= $\sigma_I \sigma_J$ (since $r = 1.0$ )
(10) = area  OBEF (since $\cos 0^\circ = 1$ )		= $\sigma_I^2$ the variance
(11) $OC =$ $\sqrt{(OA)^2 + (OB)^2 + 2(OA)(OB) \cos \theta}$		$\sigma_{(I+J)} = \sqrt{\sigma_I^2 + \sigma_J^2 + 2\sigma_I \sigma_J r_{IJ}}$
(12) $BA =$ $\sqrt{(OA)^2 + (OB)^2 - 2(OA)(OB) \cos \theta}$		$\sigma_{(I-J)} = \sqrt{\sigma_I^2 + \sigma_J^2 - 2\sigma_I \sigma_J r_{IJ}}$

## V. NOTES

## NOTE TO CHAPTER II

The formal exposition of S-theory as a system of symbolized generalizations may be supplemented by a biographic sketch of how the theory grew by laborious trial and error on statistical tables and graphs interspersed with sudden insights and revision of symbols. The presentation in this volume is misleading in exposing the theory in this chapter and then exploring it in detail, largely deductively, in the succeeding chapters for the sake of clarity and brevity of exposition. But actually the derivation of the theory was an arduous process consuming five years of inducing hypotheses from data, trying them out and discarding most of them, adopting a symbol here, a ruling defining it more sharply there, making a further checking trial of hundreds of graphs and tabulations of societal data, and so on to the gradual development of the still unfinished S-theory which is communicated here.

The starting point was a growing dissatisfaction with the verbalistic nebulousness of many of the textbooks on Sociology. Though aspiring to present a science, they seemed to move more in literary, ethical, and philosophical realms. When adorned with the results of statistical studies these usually presented some facts, some local and restricted generalizations, but little of fundamental theory towards integrating the science of Sociology into a unified system of knowledge.

Statistical studies in Sociology seldom seemed inspired by the attempt to prove or disprove, to discover or qualify, basic theory. The very possibility that statistical studies could bear so crucially on theory seemed to be considered by qualitative sociologists as forever unlikely, and by quantitative sociologists as not likely at present.

From these premises the author decided in January, 1934, while on leave in the States, and after his research in *A Controlled Experiment on Rural Hygiene in Syria*, to select a small part of basic theory in Sociology and concentrate on an attempt to build a measuring instrument adequate to test that segment of theory. The theory of social processes was selected. Most textbooks assert some set of processes and arrange them in some order, such as subtypes of association—dissociation (Von Wiese, Lumley, 2nd edition, Ref. 41) or as a series implying a continuum from conflict through competition, and through accommodation, to assimilation (Park and Burgess, Ref. 54). Why should there be so little agreement on the basic processes? How much of the difference was in different word-names for much the same referents? Is there any continuum as implied, and, if so, which processes were on it and which diverging from it? For these and many other questions the first need seemed to be operational definitions of the processes, i.e., a schedule card of behavior items with instructions for filling it out, which would thus standardize the meaning of each sociological process. Specifications of such a schedule card or scale were, therefore, drawn up, specifying that samples of behavior and its cultural products be gathered from the institutions of business, the school, the church, the family, the nation, so as to define, if possible, "conflict" in general, "competition" in general, etc. These specifications for an intercultural scale to measure and thereby define the social processes were circulated among a few colleagues for critical overhauling before seeking co-operating investigators or financing.

As a working hypothesis Park's and Burgess' list of processes was selected and a study of the literature begun as a background for inventing and selecting items for the scale. (The Social Science Research Council's subcommittee on Competition and Cooperation had not as yet published its bibliography and survey (Ref. 47) of the field, but Messrs. May, Allport, and Murphy invited the author to some of their sessions while in Cambridge.) Study of the sociological literature on "processes" led to their tentative definitions substantially as given here in Chapter X, although the notation was very clumsy and later had to be completely revised (and increased in generality) several times. The recurrence of "desideratum," "desire," and "population" concepts in these definitions led to interrelating these three concepts in the tension theory equation, cf. Chapters V and X, first in its simple scriptless form, then inventing script by script as required by the situations to which it was applied.

The theory (see Ref. 55) was then written out and its thirty-odd equations defining social processes were tabulated on one page with the factors, scripts, and some other properties compared in parallel columns. From this sheet it occurred to the author that all the process equations could be mathematically derived from the tension theory equation when the exponents varied over values of 0, 1, or 2. Accordingly the exponent was added to the simple tension theory

equation, Eq. 34, Ch. V, as in Eq. 49, Ch. V. The whole theory was revised, the processes were measured by means and standard deviations and related to the moments of frequency distributions.

The resulting classification into columns and rows on another sheet led to "predicting," hitherto unthought-of processes with such-and-such characteristics. Reflections and research through a collection of graphs of social science data and through sociological textbooks would then reveal examples of processes having those specifications. "Mobility," "dispersing," "correlating," and two thirds of the other processes outlined in Chapter X fitted themselves into the system in this way. At first, elaborate terminology was developed, coining process names as needed with the help of a large dictionary. Later the symbolism and neologisms were simplified to correspond to current concepts in Sociology as far as this could be done without misleading. Compound or aggregative processes then began popping into consciousness with their formulae ready made on waking in the morning, following reflection on them the evening before.

It was next realized (December, 1935) that many of the processes did not depend on value systems but were entirely general. They were rewritten in terms of T, I, and P units which the author had previously (1932-34) developed and published (Ref. 12) in an experimental statistical study of a societal force in primitive Syrian villages and Bedouin tents. An attempt to classify these processes as rewritten led to a crude version of what later developed into the quantific table of Chapter II. This led to a systematic thinking out of the present S-theory, using geometric concepts of vectors in  $n$ -space to aid in visualizing and to check on the internal consistency of the system.

Graphs were then decided upon (January, 1936) as the unit of observation for this research. With assistants, a loose-leaf photostat collection of fifteen hundred graphs was built up and each was expressed in a formula using T, I, P units (and some others later discarded as unnecessary). The system of rules for writing formulae was repeatedly revised, requiring the rewriting of the whole collection each time. Each ambiguity or problem as it was presented by a new graph was recorded with alternative hypotheses (mostly notational) and tested on further graphs, till one hypothesis would emerge as the best, i.e., most interconsistent, clearest, simplest, briefest, etc. The S-theory was then written up and submitted for criticism to Mr. George A. Lundberg, and finally by August, 1936, the first draft of the chapter on Change was penned.

The ambiguity in the writing of formulae was next tested by having two paid assistants write them independently with the result of three percent of discrepancy as described above.

The S-theory then went through a process of critical refinement, a process assisted by a class of Juniors at the American University of Beirut. Its presentation in the chapters of this volume was planned and a detailed 40-page outline prepared and three hundred graphs selected as a representative factual basis for the theory. In this critical review, Professor Lundberg assisted in the summer of 1937 in the Lebanon mountains. Writing the manuscript, developing its equations, checking, and proofreading were the final stages throughout 1938 and 1939.

In short, an attempt to throw light into one corner of basic sociological theory by constructing a scale designed to define in operational terms the social processes, led to the ad interim handling of the still unmeasured concepts by means of algebraic symbols for convenience. These symbols became fruitful in multiplying the sociological concepts they could express with generality and increased precision. They proved derivable from, and therefore summarizable in, the single matrix equation of S-theory (Eq. 50, Ch. II). Once this was induced, the concepts and relations between them were deduced by systematic varying of its variables. Of the large number of resulting formulae only some seven hundred were selected for publication here by the criterion that every quantitatively recorded situation in the literature of the social sciences should have a pigeon-hole, i.e., an S-formula, to fit it, but that further formulae need not be developed in advance of the publication of sets of facts corresponding to such formulae. A theory must fit all the published facts.

---

1. For a defense of thus adopting a situation for the unit of gross societal observation, see Lundberg's discussion of a sociological field of force treated as a closed system in his companion volume, *Foundations of Sociology*, in chapter 12 and elsewhere.

2. For an operational definition of "time" and "space," see Note 1 to Ch. VI of Lundberg's companion volume, *Foundations of Sociology*. For illustrations of time and space, see the three hundred S-situations appended to the chapters of the present volume, which include the symbols T or L in their descriptive formulae.

3. It is possible to use three, or five, or other sets of indices, provided they are so defined as to include the entire societal field without overlap. Thus for many studies, space can be neglected. On the other hand, for example in economic studies, monetary indices are so basic as to justify denoting them by a separate symbol and using five kinds of indices. Such a monetary index is still a subclass of characteristic as defined here, however, and is simply labeled more explicitly for convenience without affecting the logic of our analysis into four basic types of indices.

The indices of time and space have always been recognized as basic in the history of philosophy and science. Special sciences have selected other additional indices as cardinal and have thereby marked off their field as Physics did, in adding mass to space and time. Economics analyzes production into the factors of land, labor, capital (and sometimes management), all operating in time, of course. These factors in S-terms are those of time, space, population, and two characteristics peculiar to that field. Statistical textbooks trace statistical series to variations in time, space, quality, and quantity. Our attribute hypothesis (Ch. III) combines quality and quantity as subvarieties of characteristics. The population index is added to *delimit* the sociological field, for the only major component common to all the social sciences and shared by none of the sub-human fields is the human population. The concepts of S-theory may readily be extended to all biology and other non-human fields, if desired, by defining

suitable "characteristics" to be specially symbolized. But the field of this volume is human society.

4. An operational criterion of what is unitary is given later.

5. The limitation to quantitative phenomena may become progressively removable. Already our "attribute hypothesis" (see below) extends the S-theory to a large segment of qualitative phenomena. But, since the limits of this segment have not been discovered yet, its unqualified inclusion at present may prove later on to be an overclaim.

6. Any item in these six rows can be derived by multiplying the item below it by (I), or by dividing the item above it by (I). This little tabulation may help the reader whose high school algebra has grown rusty to revive his understanding of the meaning of the exponent, which is an operational symbol directing the operator as to the number of times to take the base as a factor. The exponent rule that in multiplying two indices (or any variables) their exponents are added (i.e.,  $I \times I = I^{+1+1} = I^{+2} = I^2$ ), and that in dividing one index by another their exponents are subtracted (i.e.,  $I/I = I^{+1} \times I^{-1} = I^{+1-1} = I^0 = 1$ ) may be readily verified by the reader for himself and applied in the ensuing discussion. Thus fractional exponents such as a square root can be worked out from this principle: ( $\sqrt{I} = I^{\frac{1}{2}} = I^{+.5}$ , since  $I^{+.5} \times I^{+.5} = I^{+1} = I$ ).

7. For reference, the seven hundred odd formulae of this volume, most of which are presented as equations defining a concept or stating a relation, will be numbered as equations consecutively within each chapter and abbreviated to "Eq. 1, Ch. II," etc. As every formula is an S, as defined above, every one may be set equal to S making an explicit equation of it (if it is not already written as an equation).

8. The remaining detailed formulae of spatial exponents are:

$L^0$  = specified dimensionless points in space (Eq. 2b, Ch. II)

(I): $L^0$  = (I) = a non-spatial index, i.e., an index with no spatial dimensions and no points corresponding to it (Eq. 2c, Ch. II)

$L^{+1}$  = a line in space (Eq. 2d, Ch. II)

(I) $L^{-1}$  = an index per unit of length (Eq. 2e, Ch. II)

$L^2$  = an area (Eq. 2f, Ch. II)

(I) $L^{-3}$  = an index per unit of volume (as per liter, etc.) (Eq. 2g, Ch. II)

(For examples of sociological situations involving these spatial exponents, see S. 8 and 9, Ch. II, and any S's in Ch. VIII.)

9. A process, as shown more fully in Ch. X, is but a specified form of relatively continuous change.

10. To distinguish the two T's requires further scripts to be expounded below. For more complicated illustrations of similar societal situations, see:

S. 5 and 7, Ch. IX in connection with Eq. 6, Ch. 2.

S. 8, Ch. II; S. 59, Ch. X; S. 39, Ch. XI in connection with Eq. 7, Ch. 2.

S. 22, Ch. II; S. 24, 26, Ch. X; S. 34, Ch. XI in connection with Eq. 8, Ch. 2.

S. 18, Ch. II; S. 3 and 6, Ch. X in connection with Eq. 9, Ch. 2.

S. 2 and 26, Ch. II; S. 75, 77, 80, Ch. X in connection with Eq. 10, Ch. 2.

11. For more complicated examples of populationless situations, see any S appended to Ch. III.

For more complicated examples of plurels, see any S appended to Ch. IV, and S. 1 and 2, Ch. II.

For more complicated examples of groups, see any S appended to Ch. VII, and S. 12, Ch. II.

12. These three categories can be shown to parallel in more mathematical language the categories of quality, quantity, and relation in Kant's *Critique of Pure Reason*. (See Appendix IV.)

13. Indicants to the third power, as in measures of skewness, and to higher powers as in measures of kurtosis, occur, but are still rare in social science literature.

14. First-degree equations are called linear; second-degree ones are quadratics; third-degree ones, cubics; fourth-degree ones, quartics; fifth-degree ones, quintics; etc. The degree of the S is determined by the sum of the exponents on all the indices. Thus a correlation scattergram has the quantic formula  $I^2P^1$  and is, therefore, a cubic. According to the number of different sectors, i.e., types of indices, in a situation the S may be called:

binary, if it has indices from 2 sectors as  $IP$  or  $PL^{-2}$

ternary, if it has indices from 3 sectors as  $T^{-1}IP$

quaternary, if it has indices from 4 sectors as  $T^{-1}IL^2P$

15. In this four-digit number the complement of ten is substituted for negative exponents for convenience, as in the characteristic part of a logarithm. Thus the quantic formula  $T^{-2}I^{-1}P^{+1}$  of S. 31 has 8;1;0;1, as its quantic number. Noting two special cases may prevent confusion in studying the quantic numbers of the graphs. Wherever an index appears with a plus and a minus exponent (see S. 2, 3, 26, 29) they are both recorded. Thus in  $T^{-1}T^{+1}$ ,  $|^* = 91;0;0;0$ ; or in  $T^{-1}P^{+1}P^{-1}$ ,  $|^* = 9;0;0;19$ . Thus the odd digits denote varieties of primary indices, while the even digits denote varieties of secondary indices (See Ch. III for details).

16. Any attempt to represent four variables ( $|^*$ ,  $|^1$ ,  $|^1$ ,  $|^p$ ) on a two-dimensional page is unsatisfactory. The quantic solid (a three-dimensional model, see S. 33), described later on in the geometric section of this chapter, is a more satisfactory representation.

17. The vertical bar serves as a base for scripts to show which of the four possible corners that script occupies, and therefore, to identify the script as an exponent,  $|^*$ , class script,  $|_*$ , class-interval script,  $|_*$ , or case script,  $|^*$ . It can also be thought of as unity for the class-interval script, so that the product of the symbol  $|_*$ , and a homosectoral index ( $I'$ ) gives a scale, or extension, of class-intervals,  $|_*' \times (I') = (I'_*)$  in number, each of ( $I'$ ) units.

$|_*' \times (I') = (I'_*)$  a homosectoral index in class-intervals

(Eq. 28, Ch. II)

18. Attributes (qualitative indicators) may have qualitative subclasses denoted by the class script, but can have no class-intervals as these subdivisions are of quantity.

## 19. For common examples:

σI denotes an indicator in sigma units

\$I denotes an indicator in dollar units

%I denotes an indicator in percentage units, i.e., an index number

20. In the captions defining the descripts in this formula the words, "the aggregation of," are understood, e.g., aggregation of cases =  $\sigma$ ], etc. If the value of the descript is zero, the index is non-existent. It then has the value of zero and is called a "nul" index. When no script is written on an index, a class script of unity and an exponent of unity are understood.

21. See S. 26, Ch. II; S. 2, 3, 5, 6, 7, 8, 9, 12, 13, 14, Ch. IV, for further examples. If the territory, rather than the people, were denoted in Eq. 39, Ch. II as by a map, the formula would be:  $L_V^2 : m : n$ . For a classification of attributes in a hierarchy note the formula  $I^0 : i : j : k$ . The pattern of class scripts in a hierarchy of classes or aggregations in one sector is symbolized by  $|_s$  and this will be referred to as the "classification script":

$$|_s' = |_{t:u:v}, \text{ or } |_{i:j:k}, \text{ or } |_{1:m:n}, \text{ or } |_{p:q:r} \quad (\text{Eq. 40, Ch. II})$$

the classification  
script

22. This colon symbol seems to have much in common as a statistical analogue to the logical sign for "implication,"  $\supset$ , used by Russell and Whitehead in their *Principia Mathematica*. It serves for the discontinuous data of statistics somewhat as the sign for function,  $f$ , in "y is a function of x,"

$$y = f(x) \quad (\text{Eq. 44, Ch. II})$$

But the colon asserts that the variable after the colon is a function of, or dependent upon, the variable before the colon. The function denoted by the colon, however, is not reversible, it is one-way dependence. It thus usefully differentiates the one-way implication of a correlation ratio from the mutual implication of the correlation coefficient, and societal contact from interaction in the interrelation matrix.

The meaning of the colon in specifying the matrix is a variation from the double script (denoting row and column) of conventional matrix algebra. The variation gives greater flexibility in handling matrices of all orders, ranging from the single index which is a matrix of zero degree, upwards indefinitely to include matrices of the tenth or higher degrees, such as are involved in a whole volume of census statistics. The colon separating the two scripts of the double-script notation permits specifying that the row classification and the column classification may be in different sectors (as in S. 19, 20, 28, etc.) or of one sector, but of components with different exponents (as in S. 9, 26, 29). The colon permits distinguishing matrices composed of dependent and independent variables. The double-script matrix notation connotes that the data have been rectangularly tabulated, while the colon is more general in denoting graphs, maps, or paragraph statements where the rectangular arrangement is only potential (and often imperfect with empty cells). (See S. 9, 24, 34 for matrices in non-rectangular arrangement, and S. 15 and 30 for matrices with empty or zero-



- $\neq$  denotes "is unequal to," or, "is not equivalent to"
- $<$  denotes "is less than"
- $>$  denotes "is greater than"
- $\Sigma$  denotes "the sum of"
- $\overline{\phantom{x}}$ , overlining, denotes a vector, i.e., the quantity in geometric terms
- $\begin{vmatrix} \phantom{x} \\ \phantom{x} \\ \phantom{x} \end{vmatrix}$  denotes a matrix, a rectangular arrangement of entities into rows and columns
- () , parenthesis, denotes treatment of its contents as a unit, hence denotes an "index" in S-theory
- ?, question mark, questions an assertion or an equation, labeling it as an hypothesis
- $\underline{\phantom{x}}$ , underlining, denotes indefiniteness, i.e., the existence, but not the amount, of some quantity is asserted. Thus a per capita ratio may be stated with both the numerator and the denominator left indefinite or unstated. Usually the underlined quantity has been measured but is not recorded in the situation as presented.
- ;, the semicolon, denotes that the letters it separates are combined in some unknown or unstated function, i.e., it denotes *any* of the operators without specifying which operator.

The last two symbols ( $\underline{\phantom{x}}$  and  $;$ ) with their denotation as above are newly introduced in S-theory and are not simply extensions in meaning of conventional symbols.

25. A mnemonic phrase with the initial letters in this sequence and which suggests the content of the theory is "Timed Indicators of Land and People."

26. A warning is needed here lest some students of S-theory get the notion that expressing data in algebraic symbols implies that the precision of the data is thereby increased. Thus several colleagues on reading a six-page abstract of S-theory have written the author, fearing, for example, that expressing the ordinal steps of an attitude test in algebraic symbols will imply that the units of attitude have thereby become cardinal units. On the contrary, S-notation in reserving the class-interval script for cardinal units and the point script for ordinal units should help to distinguish data, or measurements, of varying degrees of precision, and prevent the very confusion these colleagues rightly deplore.

27. The analysts were the author and Miss Angela Jurdak who wrote her M.A. thesis on *An Experimental Determination of the Reliability of S-Theory* (Department of Sociology, American University of Beirut, June, 1938). This pioneering study in experimentally measuring the objectivity of a theory in Sociology may be secured in bibliofilm from the American Documentation Institute, Offices of Science Service, 2101 Constitution Avenue, Washington, D.C. Document # 1157 by remitting ninety cents for microfilm form.

28. A third corroborating but imperfect set of evidence was an earlier test with faulty technic conducted in the summer of 1936 when the theory and the notation were constantly being revised with the help of two assistants who were medical students with no statistical training. Offsetting these negative handi-

caps to agreement was the fact that most of the first five hundred situations had been discussed together in the earlier stages of trying to standardize notation. The trend is to reduce discrepancies as the analyzers acquire practice in analysis. The discrepancies were:

1st hundred graphs.....	7.0%	8th hundred graphs.....	1.3%
2nd " " .....	4.0%	9th " " .....	2.0%
3rd " " .....	4.1%	10th " " .....	1.6%
4th " " .....	5.0%		
5th " " .....	2.3%	Total.....	30.3%
6th " " .....	1.4%	Average.....	3.0%
7th " " .....	1.6%		

A published misprint may be corrected here. The standard error of the 93% of agreement on the descriptive formula was  $\pm 1.14\%$  as printed here ( $= \pm .77 = PE\%$  as on p. 43 of Miss Jurdak's thesis) and is not  $\pm .45$  as printed in line 18, p. 630 of the author's article "A System of Operationally Defined Concepts for Sociology" Amer. Soc. Rev., Oct. 1939, No. 5, Vol. 4. The formula is

$$\sigma\% = \sqrt{\frac{\% (100 - \%) }{N}} = \sqrt{\frac{93 \times 7}{500}} = 1.14.$$

29. It is to be expected that if the experiment were repeated by persons not trained by the author but trained only by reading the exposition in this volume, the percentages of agreement might be different. How different, and whether higher or lower, such experiments themselves will say.

For further details on this experiment, the reader is referred to Miss Jurdak's thesis.

30. A function of this physical model of the quantic classification is to aid in visualizing and testing our "unit situation hypothesis." This hypothesis is that any quantitative societal situation is a unit when, and only when, its quantic formula specifies *adjacent* cells of the quantic solid. It is a "composite" unit situation, if more than one adjacent cell is involved. But if two non-adjacent cells, i.e., two non-contiguous regions, are specified by the quantic formula, then there are two unit situations, with two quantic formulae. This provides an objective definition of what is the *unit* societal situation. Thus, a combined map ( $L^2$ ) and table not involving space at all ( $L^0$ ) represent two situations, since  $L^0$  and  $L^2$  are non-adjacent arrays, being separated by  $L^1$ . Quantitative presentations involving simultaneously interrelations of persons ( $P^2$ ) and no persons ( $P^0$ ) are two situations, whereas interrelations and some characteristics of those persons might be one composite situation, since the  $P^2$  array is adjacent to the  $P^{+1}$  array.

A static and an accelerated situation,  $T^0$  and  $T^{-2}$ , are two separate situations, as are also qualities,  $I^0$ , and correlations of quantities,  $I^2$ . This hypothesis is purely a definitional one—a proposal to conventionalize what is to be called a unit situation. Its verification is not a matter of truth but purely a question of utility. It is closely related to the new field of "topology." In fact the "unit situation" is a topological definition of a single closed region in the quantic solid.

31. The journals scanned were: *American Journal of Sociology*, *American Sociological Review*, *Revue Sociologique*, *Social Forces*, *Sociological Review*, *Sociology and Social Research*, *Survey Graphic*, and several volumes of other social science journals, including *Journal of the American Statistical Association*.

32. An early form of the theory included only T, I, and P indices and only the exponent, date script, plurel script, and indicator script:

$$S = {}_p^t(T,I,P)_i^s$$

(in different notation) and was extended, by the necessity of describing all the situations that were encountered, to its present more inclusive and flexible set of scripts.

33. As such modifications cumulated, all the situations analyzed up to date had to be reanalyzed in the light of the revised notation or rulings. The procedure in adopting a modification was to invent alternative hypotheses as to notation or definitions which would include the baffling situation. These alternatives would then be tried out on more situations until the alternative that was most parsimonious, most objective, or simplest emerged. It was then applied uniformly to all the situations involving that issue. The magnitude of this task of fitting and refitting the symbolic concepts to societal data may be judged from its having consumed roughly some three thousand hours of the author's time in research.

34. "I regard this work as one of the finest sociological monographs which has appeared up to the present time. In its own particular field, namely, the technic of measuring social change, it is beyond question superior to anything heretofore published. To one who has had occasion recently to declare that the entire literature of social surveys, including his own contributions to that subject, is of very minor scientific importance, it is a great pleasure to come upon in this unpretentious volume a model of what a scientific piece of work in this field should be. The full merits of the book can be realized only through careful study of it. But I shall call attention to the two main grounds upon which I accord it the above high estimate, namely, (1) the rigor of its scientific procedure and results, and (2) its contribution to a phase of fundamental sociological theory. . . .

"To those who regard the field of social measurement as but a naive and passing aberration, even the high quality of the work under review will seem of little sociological importance. To those who believe that a large part of future sociological research must consist of a painstaking scientific check of the generalizations which today fill the Introductions to Sociology, Dodd's book will be an encouragement. It may be that a hundred years from now, Sociology will still be nothing more than a formidable collection of droll stories, impressionistic generalizations, and windy dialectics. On the other hand, it may be that we shall then have a substantial set of generalizations the probability of which, under given conditions, can be stated with some mathematical accuracy and reliability, as must be the case in every true science. Professor Dodd's volume proceeds on the latter assumption. It demonstrates a type of scholarship which does not rest entirely upon the ponderousness and antiquity of his bibliography,

and which is as yet not very common in Sociology. Nor will the customary charge that studies of this kind neglect the theoretical side of it apply to Dodd's work. 'The Theory of the Measurement of Social Forces,' submitted in the concluding part (Part IV) is one of the ablest and most stimulating treatises that has yet appeared on this subject. Briefly, there is here submitted a logical theory of the reduction of the concept of social forces to an equation based on measured entities."

35. "The most thoroughgoing illustration of the experimental method of sociology is the recent report of S. C. Dodd of an experiment in Rural Hygiene in Syria. After some years of painstaking research in preparation and testing, he developed a scale to measure personal hygiene in terms of the behavior of native Syrians. Two samples were selected for study: an experimental village in which a hygiene program was to be put on; and equated controlled villages, without such a program. Before the program was put on in any of the villages, all were measured with the hygiene scale. After two years all the villages were again measured. Differences in the second scale position of the two samples were taken as a measure of the effects of the program, other things being as nearly as possible equal in terms of the precautions taken. This really important study is not widely known. Most reviewers of the book display an astonishing ignorance of scientific procedures. In fact, their reviews are little more than so many examples of indecent intellectual exposure."

36. "A brilliant beginning has been made by Dodd, who defines a social force as that which accelerates a social change in a population. He constructed scales to describe in quantitative units the hygienic conditions of certain populations. Over a period of two years the hygienic habits of these populations changed. The degree of improvement was measured by the number of units gained in hygienic score. The unit was a 'stim,' or one person stimulated to change one unit of hygienic score per year. Thus a social force was measured in units of 'stims.' Dodd also developed a formula to compute acceleration and social momentum, and elaborated a rational theory of social forces on this basis. We agree with Dodd's principle that social forces may be measured by noting the difference between readings on some index of social position *before* a social program is applied to a particular group at a particular place and time, and the readings on the same index *after* the social program has been applied."

37. No other published references to this theory are known to the author, though the book, *A Controlled Experiment on Rural Hygiene in Syria*, has been reviewed, favorably and adversely, by:

Young, Erle F., in *Sociology and Social Research*, Vol. XX, No. 1, Sept., 1935.  
 Embree, Edwin R., in the *American Journal of Sociology*, Vol. XI, No. 3, Nov., 1934.  
 Garle, H. E., in the *Journal of the Royal Central Asian Society*, Vol. XXI, Part III, July, 1934.  
 Wiehl, Dorothy G., in the *Annals of the American Academy of Political and Social Sciences*, Vol. 182, Nov., 1935 (p. 231); in the *Journal of Educational Sociology*, Dec., 1935; in the *Revue Internationale de Sociologie*, Nos. III-IV, Mars-Avril, 1936.

38. 1. *Observed and recorded data:*

S<sub>14</sub> Equation of an analogies type of intelligence test item defining an I<sup>0</sup>-unit, a unit-attribute.

S<sub>15</sub> Equation defining a schedule card—a punched card of a census—a type of standardized case record.

S<sub>16</sub> Equation of an item of qualitative evolution—the history of the wheel.

S<sub>17</sub> Equation of a distribution of a characteristic among plurels—life insurance by nations.

S<sub>18</sub> Equation of indices presented in a prose paragraph—mortality rates by types of transportation.

## 2. Systematized data:

S<sub>19</sub> Equation of a classification—social institutions and their type parts.

S<sub>20</sub> Equation of tabulating—economic values by years and States, also, a derived equation defining an index of competition.

S<sub>21</sub> Equation defining a social force.

S<sub>22</sub> Equation of a budget balancing income and expenditure.

S<sub>23</sub> Equation of a factorial matrix from intercorrelations.

S<sub>24</sub> Equation of interrelation of employers and employees—a cartoon dramatizing a generalization.

## 3. Induced principles:

S<sub>25</sub> Equation (calculative) expanding an hypothesis—relation of capital and labor to production.

S<sub>26</sub> Equation stating an hypothesis—"population pyramids are normally isosceles."

S<sub>27</sub> Equation defining a theory—a tension theory of societal action.

S<sub>28</sub> Equation (calculative) of a law—the normal probability curve in the case of the rate of cultural diffusion.

S<sub>29</sub> Equation based on laws of probability—life insurance premium by age and duration.

## 4. Verification:

S<sub>30</sub> Equation of the outcome of a controlled experiment on communities.

S<sub>31</sub> Equation of correspondence between a national plan and its achievement—Russia.

S<sub>32</sub> Equation of evidence disproving the early Malthusian theory.

S<sub>33, 34, 35</sub> are geometric interpretations of S-theory.

39. The nine vectors which occupy a nine-dimensional orthogonal space can be only imperfectly diagramed on a two-dimensional plane.

40. Overlining denotes vector quantities; non-overlining, scalar quantities.

41. The parallelogram is determined by one vector,  $\overline{OB}$ , and the normal,  $\overline{OF}$ , to the other vector,  $\overline{OA}$ .

42. A paper entitled "A Conceptual Scheme of Society" (Amer. Jour. Soc., Vol. XLV, No. 5, Mar. 1940) by George Devereux, proposes three basic concepts for sociology which transmute into S-theory as follows:

Devereux' "individual" = persons, P

" "motion" = time, T, and "relative position"

" "relative position" = indicators, I, and or space, L

The reliability and validity of Devereux' concepts (as also of Eubank's) have not as yet been tested by controlled experiment as have the S-concepts.

## *PART II*

### THE CHARACTERISTICS SECTOR, II

*studying situations defined by  $S = T^0; I^{0,1}; L^0; P^0$*



## Chapter III

### INDICATORS, I

#### I. THE INDICATORY SECTOR, $I_i^i$

##### A. Definition and Relation to Other Sectors

An indicator is the objectively observable sign of a characteristic. It is the material or behavioristic indication, the verifiable evidence, of the existence of the characteristic it represents. It is that upon which observers will agree. It is the conventionally accepted fact as separate from interpretation or inference of the meaning of the fact. It may be a material fact, such as the number of telephones in a region, indicating the amount of telephonic communication characteristic of that region. It may be some readily observable aspect of behavior, such as the attendance at some gathering serving as an indicator of the widespread interest shown in the purpose of the gathering which characterizes that population. It may be an artificially objectified indication of subjective feelings, attitudes, or opinions, as in bringing these intangible internal states to spoken or written expression in a vote or an attitude test. It may be very obvious wherever it is directly and objectively observable, as in determining the number of sheep in a flock. Our term "characteristic," includes characteristics of individuals, of their environments, and *relations* between individuals or between individuals and their environments. A more complete specification of the meaning of the indicator, I, is found in the three hundred graphed situations in this volume in which I occurs in the formulae.

The characteristics of people and of their environment are so all-inclusive and, as yet, so often dimly perceived by human beings as to be best definable as the residuum of all societal data after the objectively-determined aspects of time, physical space, and the number of persons have been abstracted out. In the broadest sense, the total of all characteristics includes all knowledge of, or pertaining to, human beings. The vocabulary in the largest dictionaries may be considered as a suggestion of the ex-

tent of this knowledge, since the larger part of communicable human knowledge has been reduced to words. But human knowledge is so vast that, while it may be studied as a whole by the epistemologist, it must also be subdivided, in order that human beings with finite minds and limited spans of attention, may develop knowledge by segments, which are called the sciences. Each science confines itself to some characteristics, neglecting the rest, while interstitial and synthesizing sciences attempt to integrate these segments.

Accordingly, the segment that is general Sociology will be defined in this volume by four kinds of variables, and specified more precisely by their indices. The characteristics of time and space are fundamental as the medium in which all human life goes on. The human population delimits the societal segment more narrowly. Hence the number of persons,  $P$ , in any given situation is adopted as a third type of variable contributing to the defining of the sociological segment of human knowledge. A fourth type of variable is the residuum of all characteristics pertaining to a human population. No list of these would be water-tight enough to make the definiteness of boundaries required by a scientific classification. The definition of indicatory characteristics as, "all other than time, space, and population" at least avoids ambiguity in classification. These four types of the most general variables will be termed the temporal, spatial, populational, and indicatory sectors of societal knowledge. Items of knowledge within each sector will be specified by temporal, spatial, populational, and indicatory indices—a term to be defined more rigorously as each sector is discussed in turn.

All of the indicators in a situation compose the indicatory sector which is symbolized by:

$$I_i^1 = \text{the indicatory sector, indicators with all their scripts} \\ (\text{Eq. 1, Ch. III})^1$$

Subsectors of the indicatory sector may be used to define the different social sciences. Thus, Economics may be fairly well defined, as the subsector having as its nucleus monetary indicators of exchange values. Political science deals with indicators of governmental structure and function. Physical anthropology studies indicators of the human anatomy and physiology, espe-

cially in classifying persons into racial plurels. Cultural anthropology works with indicators of human culture, especially in primitive groups as a simpler starting point. Ethics as a positive science, in distinction to a normative study, has, as its field, indicators of values—whatever is desired by a plurel at some time.

Tangible indicators of such values are available in enormous diversity as suggested further on when classifying indicators by content. The list of social sciences may be enlarged. The point at issue is, that, these four sectors suffice to define sociology as “the science of those characteristics which are common to all societal phenomena,” and that either subsectors (or further co-ordinate sectors) may be conventionally agreed upon to define special social sciences (or cognate sciences overlapping with the biological and physical sciences, such as physical anthropology).<sup>2</sup>

### *B. Comparison with Eubank's Categories*

A comparison with the basic concepts used in one of the best attempts to systematize the field of Sociology may be of interest. Eubank in his *Concepts of Sociology* (Ref. 25) reduces all societal phenomena to four categories: “societary composition,” “causation,” “change,” and “products (culture).” (See S. 6, Ch. III.) His “composition” includes the single human being and the plurel, and is nearly synonymous with our population sector, P. His subdivision of “change” into “products,” and “relationships” would, in their non-temporal aspects, be “indicators” in our system, with physical space differentiated out for separate symbolizing. His “change” in the “action, or processes” subdivision, and the dynamic or time aspects of his “culture” involve our temporal sector, T, compounded with other sectors. His “causation” (subdivided into “societary energy or force” and “control”) is, in our system, a compound which is further analyzable (as it is in Physics also) into an acceleration ( $T^{-2}$ ) of something to a certain amount. In defining a societal force in Chapter XI as the accelerating of change of a population ( $F = T^{-2}IP$ ), this concept of “force” is reduced to more tangible and observable entities of time, an indicator of the change, and the number of persons changed. His “control” is similarly reducible to a force exerted by people on other people, a force combined with

an interrelation of population, the formula for which is developed in Chapter XI as  $T^{-2}IP^2$ .

Our system reduces his seven subcategories to four, which are both more parsimonious and more directly measurable. Accordingly, we see our system as not contradicting Eubank's system, but more as a further refinement of his (perhaps somewhat as his system rearranged and sought to make Von Wiese's system (Ref. 79) more objective). Certainly, the present author has taken more from *Concepts of Sociology* than from any other book on Sociology.

### *C. Order of Presentation of the Sectors*

Before proceeding to analyze the indicator sector by its scripts, note should be taken of the order in which the following chapters discuss the sectors and their combinations. In expounding the sectors, the quantic classification, visualized as a solid diagram in S. 33, Ch. II, is to be followed. But alternative routes for following through its cells are possible. The route here chosen is to take up each sector in turn in Parts II to V of this volume, beginning with the more general and important sector of human characteristics, I; going on to explore their combination with enumerated populations, P; and then combining the spatial and temporal sectors, L, T. In each sector its variations, as specified by the scripts will be taken up, beginning with the exponent which determines the sectoral quantic, proceeding to the class script which specifies the variety of indices in that sector, and going on to the class-interval script and the case script. In each part after discussing its sector alone, its combinations with the sectors previously discussed will be cumulatively built up. We shall thus proceed along the  $T^0 I^0 L^0 P^0$  array which is the three cells from the bottom leftmost to bottom frontmost in the quantic solid, S. 33, Ch. II. This array in Chapter III of Part II is discussed first as to qualitative indicators,  $I^0$ , then quantitative ones,  $I^1$ , and finally correlated indicators,  $I^2$ . Then, in Part III, The Population Sector, the array of cells just over the above array (in S. 33, Ch. II) is taken up, starting with its first cell of Plurels (Ch. IV,  $T^0 I^0 L^0 P^1$ ), its next cell where plurels are compounded with an indicator in Distributions (Ch. V,  $T^0 I^1 L^0 P^1$ ), and its last cell where plurels are compounded with second degree

indicators in Correlation (Ch. VI,  $T^0 I^2 L^0 P^1$ ). Chapter VII then takes up the top array of cells where people are interrelated with people,  $P^2$ , in a group in qualitative, quantitative, and correlated ways ( $T^0 I^{0,1,2} L^0 P^2$ ). The third index of geographical space,  $L^1$ , is then studied in its lineal, areal, and voluminal combinations with indicators and population (Ch. VIII,  $T^0 I^1 L^{1,2,3} P^p$ ). Finally, the introduction of the fourth index of time,  $T$ , takes in dynamic situations of Change, subdivided into Durations (Ch. IX,  $T^{+1} I^1 L^1 P^p$ ) and Processes (Ch. X,  $T^{-1} I^1 L^1 P^p$ ), and the dynamic situations of Accelerations (Ch. XI,  $T^{-2} I^1 L^1 P^p$ ) which include societal forces ( $T^{-2} I^1 L^0 P^1$ ) and societal control ( $T^{-2} I^1 L^0 P^2$ ). A summarizing chapter follows, applying the system as a whole by exploring the extent to which it aids in predicting phenomena in the social sciences.<sup>3</sup>

The point of thus outlining the order of the chapters ahead is to emphasize the rigorous adherence throughout this volume to the quantic classification on the basis of the mathematical exponents in the four sectors.

## II. THE INDICATORY EXPONENT, $|^i$

The characteristics of people and of their environment may be observed in three levels—as qualitative characteristics, as quantities or intensities of those qualitative characteristics, and as correlated characteristics. These three can be denoted by the indicatory exponent of zero, one, and two respectively.

$$|^i = 0, 1, 2 \quad (\text{Eq. 2, Ch. III})^4$$

### A. Qualitative Indicators, $I^0$ , “Attributes”

A characteristic may be a qualitative phenomenon, a unitary kind of something. It may be distinguished, named, and even dealt with, without any knowledge of its amount, intensity, or other quantitative aspect. It may be a unique complex of phenomena ranging from a relatively simple type, such as a picture in S. 5, Ch. III, to a relatively complicated type such as the institutions verbally represented in S. 2 and S. 4, Ch. III. (Other examples are given in S. 1 to 7, Ch. III.) Since it is a unitary kind of something it may be mathematically symbolized by an indicator with a zero exponent. A zero exponent reduces any variable to a constant with the value of unity:

$x^0 = 1$ , therefore  $I^0 = 1$  = the "attribute" or qualitative indicator  
(Eq. 3, Ch. III)

What the quality is, what kind of thing it is, is denoted by the class script. An indicator with a zero exponent representing a qualitative characteristic will be called an *attribute* ( $= I^0$ ). The statement that qualitative societal phenomena can be usefully represented by an indicator with a zero exponent (and suitable other scripts) will be referred to as the "attribute hypothesis."

The invention of the attribute symbol is a simple piece of notation but its usefulness may prove to be very great, for it brings within the fold of mathematical entities the whole range of itemized qualitative phenomena, usually deemed outside the reach of precise mathematical reasoning.<sup>5</sup> \* Of course no magical transformation of mystical entities into crisp mathematical equations is to be expected by merely dubbing a quality an  $I^0$ . But a useful new set of relations is opened up, as well as the possibility of much more being discovered, by means of this symbolic tool, as will become evident on reading further into the evidence here presented.

### B. Quantitative Indicators, $I^1$ , "Indicants"

Whenever qualitative phenomena are observed more closely they may be seen to vary in some quantitative way. The intensity of the quality may have degrees. Thus, almost every quality that has been named in language in the form of an adjective, or an adverb, can be compared in three degrees, grammatically called positive, comparative, and superlative. Thus, we think of spiritual qualities as most remote from mathematical expression, yet we think of this personality, that meeting, or this other service of worship, as being "more spiritual" than some other. Every such comparison with a "more" or "less" in it denotes an implicit quantity, some rough scale for judging, which we may not be able to put into words, but which we feel and act upon. Every qualitative entity in the form of a noun or a verb may be quantified by the frequency of its occurrence, by the quantity of such entities, by relating it to others in some kind of a series, or by analysis into its properties. These properties, in turn, may exhibit degrees of more or less in comparison with the properties of other

\* For Eqs. 4a-c, 5-7, and 8a-e, Ch. III, see notes at the end of the chapter.

similar entities. Thus, we compare Health Departments in degrees of efficiency; we compare our enjoyment on two different occasions, the relative beauty of people or artistic creations, the simplicity or convincingness of a theory, the ethical value of different acts, and so on through most of the supposedly purely qualitative phenomena in life. Usually the issue is not so much whether a qualitative phenomenon has quantitative aspects, but whether any objective and definite scale or units can be devised to measure those quantitative aspects. The issue in this form is a challenge to further research. By ingenuity and persistent effort, precision of observation may be increased, though to what extent, requires a lapse of time to answer. This hypothesis that any quality can be quantified, theoretically at least, even though practical and objective means are lacking, will be termed our "indicant hypothesis," for indicant is the term to be used hereafter to denote a quantitative indicator, symbolized by  $I^1$ . An indicant, definable as an indicator with an exponent of one, means a quantity of some kind of thing (the kind denoted by the class script).

An indicant may always be thought of as a product of an attribute and a "pure" indicant:

$$I^0 \times I^{+1} = I^{+1} = \text{the "indicant," or quantitative indicator} \\ (\text{Eq. 9a, Ch. III})$$

Multiplying a quantity by unity leaves its value unaltered, but if one particular quality is identified by the attribute, as when it has a descript,  $I^0$ , then the product is a certain number of units,  $I^{+1}$ , of that particular kind:

$$I^0 \times I^1_0 = I^1 \quad (\text{Eq. 9b, Ch. III})$$

Ordinarily such a product of an attribute and an indicant is not explicitly written but is understood, just as every algebraic letter is always understood to have a coefficient of one, if no other coefficient is written. The student should, however, realize that the symbol  $I^1$  denoting the number of units,  $I$ , of the kind specified by the singular class script,  $|$ , is decomposable into a particular attribute,  $I^0$ , times a pure number (a pure indicant  $I^{+1}_0$ ), which, when alone, simply states a number of units without stating

of what kind. The unwritten attribute which is understood from the descripts to be present, is termed an *implicit* attribute.<sup>6</sup>

This product of a qualitative indicator and a pure quantitative indicator is fundamental in dealing with the other indices of population, space, and time. (Compare Eq. 4, Ch. IV; Eq. 5, Ch. VIII; and Eq. 8, Ch. IX.)

The units of an indicant may be called I-units. This term is convenient for this technical denotation, as a general class concept for the units of all kinds of characteristics is much needed in societal statistics where various terms, such as "scores," "points," "credits," "percentage units," "index units," and other specialized names of units, all have restricted connotation.

An index may next be defined as a function involving an indicator. It may be a sum of indicators, or a product of indicators, and time, space, population, or compounds of these. Products and sums include ratios and differences as the product of a reciprocal and as the sum of positive and negative quantities. The index will be symbolized by an indicator in parenthesis, (I).<sup>7</sup> \* Wherever desired, an index may be written explicitly, by specifying its factors, or its addends, within the parenthesis (such as  $(IP^{-1})$  denoting a mean or per capita indicator). An index is a compound indicator because it has the properties of an indicator and enters into further compounds with other indices.

An important class of index is the dynamic one of an indicant per period (a velocity of change or process). This dynamic index is symbolized as  $(IT^{-1})$ . Its units are events, acts, or things happening in time. Examples are births per year, wages, or expenditures per week, tons per month, telephone calls per day, etc. (Fuller examples and definitions are reserved for Chapter X, dealing with Change  $(T^{-1})$ .)

### C. Correlated Indicators, $I^2$ , "Correlates"

The indicator with an exponent of plus two ( $I^{+2}$ ) occurs in the correlation of two indicators. In essence it is the cross-classification of all the classes of one indicator with all the classes of the other indicator. Since this is usually combined with the frequency of a third index (such as persons or dates) in the form of a contingency table in the case of attributes, and in the form of a

\* For Eq. 10, Ch. III, see notes at the end of the chapter.

correlation scattergram in the case of either indicants or indices, discussion of it is deferred until Chapter VI on Correlation, ( $I^2$ ). The indicator with an exponent of 3 occurs in measuring the skewness of a distribution and in other functions of the third moment. This and higher exponents are studied in Chapter V as they occur only as compounds with other indices. In the special case where the two indicants are identical, instead of the correlation coefficient, we have the variance of the indicant.

Turning from the positive indicatory exponents to the negative exponents, the most common type is the simple index number, or ratio of one indicant to another. The divisor is usually a particular or constant value of the variable numerator, such as in index numbers relating the annual amounts to the amounts in some year taken as a base. This index number (not to be confused with our more inclusive "index") simply changes the units in which the numerator indicant is expressed. Its observed units are changed to units relative to the denominator. The indicator with a negative exponent never occurs alone. It is a divisor and always requires a dividend.<sup>8 \*</sup>

An indicatory exponent of minus two is rare. It would denote an index number recast into units of still another indicant, a kind of index number of an index number.

The meaning of the indicatory exponents may be summarized as follows:

$$\begin{array}{l}
 \left. \begin{array}{l} I^0 \\ I^{+1} \\ I^2 \\ I^{-1} \\ I^{+1-1} \end{array} \right\} \begin{array}{l} = \text{an attribute (Eq. 12a, Ch. III)} \\ \text{a qualitative characteristic} \\ = \text{an indicant (Eq. 12b, Ch. III)} \\ \text{a quantitative characteristic,} \\ \text{I-units} \\ = \text{correlated indicators, 2 cross-} \\ \text{classified characteristics} \\ = \text{something per unit of the indi-} \\ \text{cant} \\ = \text{an index number} \end{array} \quad \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{an ordinary indi-} \\ \text{cator, a qualita-} \\ \text{tive or a quanti-} \\ \text{tative characteris-} \\ \text{tic} \end{array} \\
 \begin{array}{l} I^0 = \\ \text{an indicator} \\ \text{to any power} \end{array} \quad \left. \begin{array}{l} I^{+2} \\ I^{-1} \\ I^{+1-1} \end{array} \right\} \begin{array}{l} \\ \\ \end{array} \quad \begin{array}{l} \text{(Eq. 12c, Ch. III)} \\ \text{(Eq. 12d, Ch. III)} \\ \text{(Eq. 12e, Ch. III)} \end{array} \\
 \\
 (I) = \sum_1^z I^i (I')^0 = \text{an index, a compound indi-} \\
 \text{cator; the sum and/or prod-} \\
 \text{uct of indicators and indices} \\
 \text{from other sectors.} \quad \text{(Eq. 12f, Ch. III)}
 \end{array}$$

\* For Eq. 11, Ch. III, see notes at the end of the chapter.

III. THE INDICATORY CLASS SCRIPT,  $I_i$ 

## A. Definitions

The class script on the indicator,  $I_i$ , denotes the number of indicatory indices. The prime on the class script  $I_i'$  (or  $I_i$ ) denotes the particular *identified* indicator (i.e., one class of characteristic).<sup>9</sup> It states the qualitative kind of indicator. The lower case letter,  $i$ , denotes the aggregation of indicators of different kinds.<sup>10</sup>

In order to understand what an indicator is, the legends of the graphed situations, S. 1 to 15, Ch. III, should be studied. Usually the indicators are evident from the title and captions of the table, graph, or other presentation of quantitative data. For borderline cases, the rules in the Appendix have been induced to increase clarity of definition and reduce ambiguity between different analysts in identifying the indicators of a given situation. A class script may be expanded into a classification by subdividing the class into subclasses.

In general a first class script may be thought of as a principle, often called the "basis" of classification, which is the broadest class of the phenomena presented in the S-situation. Its qualitative subclasses are denoted by the subclass scripts  $I_{i':j:k}$ , etc. This can be expressed with more generality as any organization, hierarchical system, or classification of characteristics, either qualitative or quantitative:

$$I_{i':j:k:\dots:s} = \text{a hierarchical classification of characteristics} \\ (\text{Eq. 13, Ch. III})$$

When dynamic characteristics of a person are specified, this formula can describe personality as an organization of traits, and "human nature" as the collection of "instincts," or native human traits (if they should ever be specifiable!). This formula describes an institution when the class scripts specify its type parts, as in Chapin's analysis (S. 4, Ch. III) of institutions into the type parts of property, attitudes, a code, and symbols. It can describe a culture, subdivided into culture complexes and composed of culture traits, in general or in particular, as specified by the legend stating what the symbols represent. The form of the situation as a hierarchical classification is common to them all. Of course the S-notation performs no magic alchemy in that, if the recorded data which are presented in this form are inade-

quate, inaccurate, or inappropriate, the formula will include these faults. The formula merely requires specifying that the data be in the *form* of classes and subclasses and sub-subclasses, etc. But by naming, listing, and arranging these classes, the formula may stimulate greater exactness in specifying what is denoted by such inclusive and vaguely defined concepts as "human nature," "personality," and others.

What is a subclass and what is a class, is relative. It is defined by what is presented in the table, graph, etc., which defines a given S-situation. In general, the broadest qualitative class in that situation is a class, its subdivisions are subclasses; both together are a classification. By changing the situation, as in omitting part of a tabulation, the classification may be changed. *An indicator is exactly defined as only relative to a defined situation.* This disposes of the quest for absolute, elemental entities for sociological analysis. It avoids the opposite quagmire of verbal relativities by its exactness in defining the datum, or the situation in a table, graph, etc., relative to which the definition of the indicator can be proved experimentally to have high objectivity, i.e., agreement between different observers. This relative definition of an indicator is, mathematically, that an indicator is one array of the matrix which defines the situation; an example is the achieved freight tonnage column in the Russian 5-Year Plan as presented in the matrix, S. 31, Ch. II. In Eubank's tabulation of the categories of phenomena, S. 6, Ch. III, there are four qualitative indicators ( $I_i^0$ ) cross-classified against two others, as shown by the four columns and two major rows of the matrix, and also by the two multiple class scripts. In the graph of farm modernization (S. 8, Ch. III), the five indicators denoted by the multiple class script are the five rows tabulating five kinds of farm equipment. What is an indicator becomes very definite when stated thus. Tabulating the data into the form of matrices then is an operational technic resulting in a more definite "operational definition" of an indicator.

### B. "Dissimilarity" or "Differentiation" <sup>11</sup> \*

The size of the class descript measures the degree of "dissimilarity" or "differentiation," existing in the situation as presented.

\* For Eqs. 14a-c, Ch. III, see notes at the end of the chapter.

(The process of "dissimilarizing" and its reverse "similarizing" will be dealt with in Chapter X when considering dynamic phenomena.) The larger the number of kinds of things in a situation and the greater the hierarchy of subclasses, the more differentiated the situation is. Thus, in the graph of city noise, S. 1, Ch. III, the top half of the graph, which analyzes noise into eight major sources, is proportionately less differentiated than the whole graph, thereby differentiating these eight into sixty-seven more specific sources. For another example, S. 3, Ch. III, "Objectives of Agriculture," is the summary of a booklet expanding the six objectives into subobjectives with an exposition of each. The situation defined by the booklet obviously has greater differentiation than the situation defined by the summary page and would be so denoted by subclass scripts (such as  $I_{1:j:k}$ ).

### *C. A Classification of Indicators on the Basis of Their Content*

The vast scope of indicators of the characteristics of people and their environments has been suggested. A fuller outline attempting a tentative classification of indicators will now be presented. This outline, modified from Bernard's classification of environments (Ref. 3, p. 75), seems the best available to systematize the indicatory sector. But it falls far short of fulfilling the canons of classification in that many indicators will undoubtedly turn up which have no class clearly made for them, or which might go into more than one class. The canons of total inclusiveness and of no overlapping will be unfulfilled to that extent. Its imperfections, however, may stimulate others to improve upon it. The aim in selecting the class rubrics has been the practical one of making the class boundaries as definite and objectively determinable as possible.<sup>12</sup> The subclassifying has been carried to the point where measurable indicators of the societal characteristics seem determinable.

1. The physical (inorganic) indicators <sup>13</sup> (Physical Sciences, cf. S. 5, Ch. II)
  - a. Temperature—degrees centigrade
    1. Daily,<sup>14</sup> seasonal, and annual averages and deviations
  - b. Light—hours of sunlight, daily, seasonal, and annual
  - c. Electricity—lightning, magnetism (compass), static, cosmic rays, volts, etc.
  - d. Air
    - (1) Density—barometric pressure

- (2) Motion—kilometers per hour of breeze, wind, tornado
- (3) Purity—smoke, dust, etc., particles per c.c. See S. 18, Ch. VIII
- (4) Chemical elements—proportion of oxygen, noxious gases, etc.
- e. Water
  - (1) Gaseous—humidity—percent of saturation in the atmosphere
  - (2) Solid—snow, hail, ice—centimeters per year, thickness, frequency of bergs
  - (3) Liquid—annual rainfall, volume per period in springs, reservoirs, and rivers. See S. 40, Ch. XI
- f. Soil
  - (1) Size of particles—rock, gravel, sand, earth, and dust
  - (2) Chemical composition—proportions of the chemical elements and their compounds
- g. Subsoil—mined natural resources
  - (1) Coal, oil, iron, etc.—tons, volume, or percent in an ore, or monetary values
  - (2) Geological formation—igneous, aqueous rock, etc.
- h. Natural physical processes
  - (1) Earthquakes—seismograph record, cost in life and money
  - (2) Tides—height, shore topography
  - (3) Gravity—cms. per sec. per sec. towards the center of the earth
  - (4) Radiation—radioactivity of radium, etc.
  - (5) Combustion—area, temperature, damage in lives or money, fires, volcanic eruptions, lightning strokes, etc.
  - (6) Chemical composition and decomposition—heat units, volume or weight per period of various ingredients and products
  - (7) Erosion—area or volume per year
  - (8) Other physical processes
2. Biological or organic indicators (Biological Sciences, cf. S. 29, Ch. X)
  - a. External to the individual (Botany and Zoology)
    - (1) Micro-organisms—density per c.c. in water, blood, etc. Mortality and morbidity rates caused by germs
    - (2) Insects and parasites—density, geographic area or percent of population infested, morbidity rates, financial damage, etc.
    - (3) Larger plants used for food, clothing, shelter, etc.—area found, tons or volume collected, financial values, etc.
    - (4) Larger animals used for food, clothing, etc.—heads, pelts, tons, financial values
    - (5) Harmful larger plants and animals—numbers, areas, cost of damage in units of human life, labor, or money
    - (6) Prenatal environment—biochemical indicators, blood sugar, position of foetus, pelvic measurements of mother, percent breech presentations, etc.
  - b. Internal to the human individual (Human Anatomy and Physiology)
    - (1) Anatomical characteristics
      - (a) Colors of skin, eyes, hair
      - (b) Cephalic and other length indices
      - (c) Hair texture
      - (d) Height and weight
      - (e) Other indices of special interest to the doctor, athletic coach, physical anthropologist, beautician, etc.

- (f) Fingerprints, palm prints, ear molds, retinal photos, and other identifying patterns
- (2) General physiological processes
  - (a) Reproduction—births per specified population per period. See S. 35, Ch. VI; S. 9, Ch. XII
  - (b) Digestion—quantities of food and water consumed daily, basal metabolism rates
  - (c) Excretion—frequency, volume, consistency, color, chemical, and bacteriological analyses
  - (d) Respiration—volumes, basal metabolism rates, inspiration-expiration ratios, etc.
  - (e) Circulation of the blood—blood pressure, heart beating, color, etc.
  - (f) Muscle tonus and strength—athletic tests, electrical instruments
  - (g) Growth—age, size, weight, height, indices of abilities
  - (h) Death and decomposition—mortality per population per period
- (3) Neural—glandular processes (autonomic nervous system—Endocrinology)
  - (a) Perspiration—psycho-galvanic reflex instrumental readings
  - (b) Adrenalin secretion—blood sugar, blood pressure, temperature, flushing and blanching, saliva secretion, etc.
  - (c) Thyroid secretion—speed or sluggishness of mental and physical reactions, basal metabolism indices, thyroid size, iodine content
  - (d) Gonadal secretion—indicators of rhythms and extremes of sexual interest, as distribution of nocturnal emissions, copulations
  - (e) Other responses influenced by the pituitary, thymus, and other glands of internal secretion
- (4) Neural processes (Physiological Psychology)
  - (a) Receptors—indicators of sensory acuity
    - (i) Visual—
      - standardized reading distances
      - color discrimination
      - blindness scales
      - tests for cross-eyes, nystagmus, eye jumps in reading, etc.
    - (ii) Auditory—
      - distance of standardized sound heard
      - tests such as Seashore's for discrimination of pitch, intensity, rhythm, consonance, tonal memory
      - physical indicators of decibels, vibrations per second, beats, etc.
    - (iii) Gustatory and olfactory—
      - discrimination tests of bitter, sweet, salty, sour tastes and blends, and of spicy, flowery, fruity, resinous, foul, and scorched smells
    - (iv) Skin and other senses—thresholds and discrimination tests for pressure, pain, warmth, coolness, kinaesthetics of joints and muscles, posture
    - (v) Internal organic sensations—hunger, thirst, lust, fear, and anger symptoms indicated in speech or other behavior, or instrumentally
  - (b) Brain—
    - indicators of the functioning of the central nervous system

retentivity and aphasias

intelligence and feeble-mindedness

psychoses (insanity)

As far as inferred from available evidence to be chiefly determined by hereditary, glandular, and physiological factors, more than by instructional factors. See S. 11, Ch. V; S. 44, Ch. X; S. 2, Ch. XII

(c) Effectors—

indicators of capacity for speech and other co-ordinations, the particular expression of which is more culturally determined range, speed, strength, endurance, delicacy, etc., of motor co-ordinations that are relatively untrained <sup>15</sup>

3. The cultural (super-organic) indicators (Social Sciences)

a. Physico-cultural (material) indicators (Applied Physical Sciences, such as Engineering). See S. 18, Ch. XI

(1) Tools—number, blueprint specifications, efficiency ratings. See S. 26, Ch. X

(2) Weapons—number, effective distance, area, and force. See S. 13, Ch. II

(3) Ornaments—catalogue specifications of materials, sizes, shapes, prices

(4) Machines—(as for tools), horsepower, production per period. See S. 25, Ch. X; S. 28, Ch. XI.

(5) Transportation systems—kilometrage of roads, railways, canals, airways, wiring, piping; number of vehicles, ships, airplanes, ports; tonnage; passenger miles; capital invested; profits. See S. 23, Ch. VIII; S. 59, 60, 68, 84, 83, Ch. X; S. 39, Ch. XI

(6) Communication systems—kilometers of telephone, telegraph, and cables; telegrams and cables per period; pieces of mail per period; pages of print, sales or circulation of printed matter; broadcasting stations, hours, receivers; movie films, attendance, receipts per period; number of lectures, meetings, conferences, visits, etc.; vocabulary in dictionaries; Braille publications, and signs for the deaf; etc. See S. 72, Ch. X; S. 15, 16, 26, Ch. XI; S. 22, Ch. XII

(7) Household equipment—architects' and contractors' specifications, inventories, Chapin living-room index of socio-economic status, housing schedule cards, rooms, or floor area per person. See S. 6, Ch. II; S. 30, 73, Ch. X; S. 8, Ch. XI

(8) Equipment of office, store, factory, farm, mine, and fishery—inventories, financial values, production and other ratings. See S. 8, Ch. III.

(9) Apparatus for scientific research, religious worship, education, amusement, and other specialized purposes—inventories, costs, ratings, age. See S. 32, Ch. XI

b. Bio-cultural indicators (Applied Biological Sciences, such as Agriculture)

(1) Non-human <sup>16</sup>

(a) Useful bacteria for fermenting, inoculating, etc.

(b) Domesticated plants (and derivatives) used for food, clothing, shelter, medicines, ornaments—hectares cultivated. See S. 12, 14, Ch. VIII; S. 44, Ch. XI. Tons, volume, or number produced, monetary values

- (c) Domesticated animals (and derivatives) used for food, clothing, etc.—heads, tons, financial values
- (d) Domesticated animals used for draft and power—heads, horsepower, financial values. See S. 25, Ch. X

c. Personal-cultural indicators (Psychology)

- (1) Behavior of one person as a function of the stimulus-situation, S, organic factors, O, and previous experience, E. ( $B = S;O;E$ ). Among the many bases for classifying the behavior of an individual, a few examples, of use to sociologists, are:

- (a) Classification by time sequence. See S. 21, 31, Ch. X

- (i) Stimulus



- (ii) Cognition —Information tests. See S. 34, Ch. II



- (iii) Interest —Interests tests such as Strong's for vocations



- (iv) Desire —Indicators of emotional reactions, psychoanalytic indicators, speech and other indirect behavioristic indicators. See S. 27, Ch. II



- (v) Attitude —Attitude tests such as Thurstone's, Allport's, etc. See S. 12, Ch. II; S. 11, Ch. III; S. 8, Ch. V; S. 1-5, 12, 13, Ch. VII; S. 18, 52, 54, Ch. X; S. 24, Ch. XII

- (vi) Action

- (Response)—Verbal, photographic, and other records of direct observation (diaries, life histories, etc.). See S. 24, Ch. II



- (vii) (Values) —Realization of satisfactions, from the action, images of which were part of the stimulus-situation. See S. 13, Ch. II; S. 3, 13, Ch. III; S. 24, Ch. X; S. 31, Ch. XI

- (b) Classification by frequency

- occasional to habitual—frequency of recurrence. See S. 38, Ch. X

- (c) Classification by evaluation.<sup>17</sup> See S. 5, Ch. VII

- negative responses—indicators of pain, withdrawal, disapproval

- neutral responses—indicators of neutral responses, indifference

- positive responses—indicators of pleasure, approach, approval.

- See S. 34, Ch. XI

- (2) Behavior of a plurel of persons (Cultural anthropology in part)

- Similar classifications with appropriate terms as:

- "Interest" becomes "public opinion"—votes, editorials, telegrams, audiences

- "Habits" become "customs," "folkways"—schedule cards

- "Values" become "mores," etc.—indicators of the group's positive evaluation in terms of speech reactions, time, effort, money, sacrifice of other values, and punishments of violators, to conserve the mores at issue. See S. 12, Ch. IV; S. 3, Ch. XII

- d. Group-cultural indicators—behavior of groups (Social Psychology, Cultural Anthropology, Institutional Sociology), derivative institutionalized combinations of the foregoing, organized for purposes of social control

(1) Segmental. Cf. S. 19, Ch. II; S. 16, Ch. XII

<i>Institutional Groups</i>	<i>Interest* Controlled</i>	<i>Some Indicators Roughly Suggested</i>	<i>Examples in Graphed Situations</i>
(a) Domestic	Family	Sex, age, kinship and residence indicators	S. 6, Ch. II; S. 10, Ch. III; 11-IV; 1-V; 35-VI; 14-VII; 6-IX; 13, 37, 76-X; 43-XI; 5, 6-XII
(b) Economic	Wealth* <sup>18</sup>	Occupation and income indicators	S. 7, 10, Ch. IV; 14, 27, 32, 36, 39, 40, 45, 46, 47, 56, 69, 79-X; 25-XI; 20, 26-XII
(c) Political	Power	Party, office, police, military, and nationality indicators	7, 10-VII; 22-VIII; 10, 35, 61, 63-X; 27, 33-XI
(d) Educational	Knowledge*	Examinations, degrees, honors	34-II; 13-V; 30, 31, 32, 41-VI; 9-IX; 5, 70-X; 5, 29-XI; 4, 10-XII
(e) Medical	Health*	Certificates, diagnoses, morbidity, mortality, hygienic ratings	30-II; 12-III; 9-IV; 1, 6, 7, 9, 12, 62-X; 7, 19-XI; 13-XII
(f) Recreational	Sociability* Refreshment	Fraternal society and club membership; friends; frequency, variety, and number of participants in sports, and amusements, entertaining guests	7-V; 13-VIII; 4-IX; 55-X; 30-XI
(g) Ethico-religious	Rightness*	Scriptures of each religion, pulpit and editorial exhortations, rites and ceremonies	1-IV; 2-V; 34-VI; 10-IX; 42, 18-X; 10, 14-XI
(h) Aesthetic	Beauty*	Exhibits, museums, concerts,	24-XI

<i>Institutional Groups</i>	<i>Interest* Controlled</i>	<i>Some Indicators Roughly Suggested</i>	<i>Examples in Graphed Situations</i>
		institutes, ticket receipts, ratings	
(i) Linguistic	Communica- tion	Dictionaries, grammars, courses	40-VI; 2- VIII; 72-X; 1-XII
(j) Racial	Race	Coloring, ce- phalic indices, etc., blood typ- ing	15-II; 5-IV; 1, 2, 12-VII; 1, 11, 22-VIII; 15, 66-X
(2) Communal	common interests combining segments, e.g., Nomadic, agricultural, and urban communities in Arabia; Hindu, Moslem, English communities in India; Arab (Christian, Moslem), Jewish in Pales- tine; Italian, Jewish, Negro communities in New York City; "Gold-coast," slum communities in a metropolis; Mormon, gentile communities in Utah; Proprie- tors, serfs, or tenant peasants; Foreign minorities, natives; Summer resort club, rural residents; etc.		weighted per- cent of above indicators shared, S. 4-V  S. 3-XI

The above classification will probably not satisfy the physiologists, psychologists, and other specialists any more than it satisfies the author.<sup>19</sup> This unsatisfactoriness may, to the specialist, be partly due to an emphasis by the sociologist on factors most relevant to *societal* phenomena, which may differ from the relative emphasis on topics within a science. It is due in part to ill-defined categories, which often overlap on the one hand and fail to provide a suitable pigeonhole on the other. Also, it is partly due to the trespassing by a specialist in one field into other fields. The justification for presenting it at all is that it may give the student a somewhat orderly and comprehensive reminder of the extent to which quantitative indicators have been developed, especially in the field of the social sciences.

It may be noted that the inadequacies of the above classification do not affect the quantic classification of S-theory. All indicators are in two cells of the quantic table, namely the attribute cell for qualitative characteristics, and the indicant cell for quantitative characteristics.<sup>20</sup> \* Therefore, changes in the above list of

\* For Eqs. 15a-d, Ch. III, see notes at end of the chapter.

indicators, provided that time, space, and the number of people were not included, would not affect the precision of the larger quantic classification.

This classification of indicatory characteristics is not a definition of them. The student should not hesitate to call some phenomenon, or property of phenomena, a "characteristic" simply because he may find no class for it in the classification above. Anyone is free to call anything a characteristic in this system, as long as it is not time, space, or population. For scientific purposes, however, it is incumbent upon him to present evidence as to the limits, reliability, and validity of the indicators he proposes for that characteristic (if such indicators have not already been established).

The relativity of indicators needs stressing again, as students often desire a more absolute definition for intellectual anchorage. An indicatory, or any other homosectoral index, is a classification and may have classes and subclasses. The classification above is then a classification of classifications in an indefinitely expandable hierarchy of human knowledge. What in this hierarchy is a single index is defined by the situation, *as presented* in the table, graph, or paragraph analyzed. It may be a minute subclass, or it may be as broad as all knowledge, as in S. 5, Ch. II, the classification of the sciences.

#### IV. THE INDICATORY CLASS-INTERVAL SCRIPT, ${}_i|$

##### A. Definitions and Notation

The next script to be taken up is the class-interval script of the indicator. Its symbol is a pre-subscript. A lower case letter,  ${}_i|$ , denotes the number <sup>21</sup> of aggregated class-intervals of that index. As usual the prime on the descript,  ${}_i'|$ , identifies a particular one of the class-intervals.<sup>22</sup>

##### B. A Classification of Indicators on the Basis of Their Precision

The class-interval script ordinarily specifies the units and is, therefore, intimately involved in the whole question of what constitutes true units and measurement. Our answer to this question is that man observes phenomena with varying degrees of precision which can be roughly described as:

1. Attributes	= $I^0$	(Eq. 16a, Ch. III)
2. All-or-none indicants	= ${}^2I$ or ${}^{1.0}I$	(Eq. 16b, Ch. III)
3. Ordinal indicants	= ${}^iI$	(Eq. 16c, Ch. III)
4. Cardinal indicants	= ${}_iI$	(Eq. 16d, Ch. III)
5. Matrices of indicants	= ${}^iI_i$	(Eq. 16e, Ch. III)
6. Indices	= $(I)$	(Eq. 16f, Ch. III)
7. Calibrated indices	= ${}^{a:z}{}_iI \pm \sigma$	(Eq. 16g, Ch. III)

## 1. ATTRIBUTES

Observation starts with noting a quality, the existence of a something. To make this communicable it is named, the word-name becoming a symbol more conveniently handled than the qualitative phenomenon itself. This name may be our attribute,  $I^0$ , representing a qualitative characteristic. It may be observed further by noting the properties, subclasses, etc., of a qualitative sort.

## 2. ALL-OR-NONE INDICANTS

As soon as any quantity of the phenomenon, larger than one, is observed, it ceases to be a constant with unit value and becomes a variable. Some phenomena occur in nature in readily enumerated form, such as the number of persons, dogs, trees, etc. But another primitive form of quantifying begins when the absence of some phenomenon is noted. It is thus seen to vary over a range of two points, presence or absence. At this point, the constant unitary quality becomes a variable, a quantity of that quality.<sup>23</sup>

By assigning a value of zero to the absence of the phenomenon, and since its presence, an attribute, has already been defined as of unit value, this present-or-absent characteristic becomes the all-or-none indicant: 0, 1. The familiar percentage is always the mean of an all-or-none variable.<sup>24</sup> \* (See S. 8, Ch. III.)

## 3. ORDINAL INDICANTS

The next step is the crude distinguishing of degrees of the qualitative phenomenon. By comparing it with another, or with itself at other times, it is seen to be "more or less" in one instance than in another. Thus, a nation's prestige is said to rise or fall

\* For Eqs. 17 and 18a-c, Ch. III, see notes at end of the chapter.

after some great events; candidate A is judged to be better than candidate B, who is better than C; child A is asserted to be more lovable than child B, etc. This judgment starts a rank series which is readily extended from two entities to a comparison of more entities, N in number. Such a series of ranks are ordinal numbers, i.e., numbers with relative magnitudes stated without the assertion of equal intervals between them.<sup>25</sup> \* Since the ordinal indicant is a series of points without the connotation of equal units between points, the case (or "point") script expresses this fact. An ordinal indicant is, therefore, usually symbolized by <sup>1</sup>I. (See S. 24, Ch. II, and S. 12, Ch. V.) All comparative and superlative degrees of adjectives or adverbs are ordinal quantities.

#### 4. CARDINAL INDICANTS

The next stage in refining the precision of quantitative observation (i.e., measurement) is to discover, or define by convention, either an objectively determinable amount or one or two objectively determinable points. This amount (or region on either side of the one point or between the two points) is taken as the standard, or unit, and other amounts, by comparison, are determined as multiples or fractions of such a unit. Familiar examples are the freezing and boiling points of water determining a temperature interval, one hundredth of which is defined as a degree centigrade. The act of divorcing defines a point on a continuum of marriage adjustment. Going to bed or being absent from work defines a point on a health scale and defines morbidity rates as a percentage of the all-or-none status of being sick or not sick. Death is a point similarly defining mortality rates. The average score of children of each chronological age can define the corresponding mental age points. For indicants where the number of units is arbitrary (as when the number of items inserted by the examiner fixes the score range in some test of ability) the natural or absolute zero of raw scores is not a stable reference point. Here the mean score of a population and the standard deviation units of departure from the mean furnish a standard origin or fixed point of reference and a standard unit (<sub>o</sub>). (Its stability in other populations depends, of course, on the representativeness of the population on which it was standardized.) Amounts which are

\* For Eqs. 17 and 18a-c, Ch. III, see notes at end of the chapter.

expressed as multiples of a standard amount (i.e., equal and interchangeable units) are called cardinal numbers by mathematicians. The cardinal indicant is symbolized in S-theory by  $\{I\}$ , an indicant with equal class-intervals, or by units,  $i$  in number.

## 5. MATRICES OF INDICANTS

In societal phenomena, it is not enough to observe the quantity of some phenomenon in one place at one time and place. Often observations must be taken in series, yielding a collection of numbers for different dates, different regions, and different characteristics. Such data arranged in an orderly tabulation of rows and columns are called a matrix. Any rectangular arrangement of numbers is a matrix. Every statistical table is a matrix, and the data of every graph in this volume can be recast into a matrix form. To do so is an operational technic which often facilitates analysis by the rules of matrix algebra.<sup>26</sup> Some of the uses of matrices which sociologists could do well to develop are listed below:

- a. Their arrays are operational definitions of indices. (The whole quantic classification is a fourth degree matrix specified by indices in general.)
- b. They enable handling data of greater complexity than can otherwise be handled in orderly fashion. (See S. 43, Ch. XI.)
- c. They yield insight into relationships as in the correlation scattergrams, the interaction matrix (Ch. VII), the matrix of societal control (Ch. XI), etc.
- d. They enable analysis of phenomena into uncorrelated elements which may prove to be useful sociological "atoms." (See Ch. VI.)
- e. They often stimulate new discoveries and prediction from empty (nul) cells, trends, etc. (The formulae and new properties of "societal interaction" and "control" were thus discovered.)
- f. They enable the converting of sociological data into geometric terms where reasoning can be more rigorously checked. (See last section of this chapter.)
- g. They define terms like "culture base"; "culture complex"; "culture traits" (as a matrix, a subsection thereof, and an array respectively, of items from a schedule card).

The numbers in a matrix, as it stands, are neither added nor multiplied together. They are simply collected side by side in what is termed an "aggregation." Throughout this volume the terms "combination" and "combining" denote either an aggregation into a matrix, or a summation, or product (including differences, ratios, powers) into an index.

The term "descript" will denote the three aggregative scripts (class scripts, class-interval scripts, and case scripts) in contradistinction to the exponent denoting self-multiplication. A descript is always aggregative, except when the operational symbol of the summation sign explicitly converts it into a single number. Thus,  $I_i$  = an aggregation of indicants,  $i$  in number:

$$(\sum_1^i I) = (I_{\Sigma i}) = \text{an index, a sum of indicants} \quad (\text{Eq. 10a, Ch. III})$$

The *singular* descript, ( $i'$  |,  $i'$  |, or  $|i'$ ), a primed descript, denotes one cell of a matrix. The *multiple* descript, ( $i$  |,  $i$  |,  $|i$ ), a lower case letter, denotes an array of the matrix.<sup>27</sup> The conventional symbol of a matrix is to enclose it between double lines  $|| M ||$ . The descripts denote a matrix in S-notation.

$${}_i I_i = || I || \quad (\text{Eq. 19, Ch. III})$$

## 6. INDICES

After observations have been assembled in an orderly matrix, some sort of simplifying or summarizing is needed to bring the complex table of numbers down to fewer numbers, or to a visual pattern which is within the limited human span of attention. Graphs convert the tabulation into a visual pattern where relations are more readily comprehended and remembered. The various summarizing indices of Statistics also serve this purpose. Thus, ratios, totals, means, standard deviations, correlation coefficients, trend slopes, etc., are all indices which are operationally defined by their formulae which direct the operator as to how to compute them. Indices then are a further extension of measurement in reducing complicated and often incompletely homogeneous data to quantitative expression in a linear scale.

## 7. CALIBRATED INDICES

The final stage of precision in observation is to calibrate the measuring instruments so as to determine their degree of accuracy.

The scientist must know to what extent sampling errors, observational errors, the range of observation, inequality of units at different regions of a scale, the choice of origin or zero point, bias from the type of index formula used, and other errors of any kind may have falsified, or rendered uncertain, his measured observations. The experimental detection and isolation of the many types of errors and their elimination experimentally or statistically is a large field of methodology. A brief outline of the chief forms of calibration of societal measuring instruments may suffice here. Calibration of indices includes:

a. Limits—definition of the field of observation

(1) Temporal—the bounding dates and periods  $= {}^tT:$   
(Eq. 20a, Ch. III)

(2) Geographic—the region specified  $= {}^1L_1^2:$   
(Eq. 20b, Ch. III)

(3) Populational—the particular plurals or persons observed  $= {}^pP_p:$   
(Eq. 20c, Ch. III)

(4) Indicatory units—the definition of the standard used  $= {}_iI,$   
(Eq. 20d, Ch. III)

(5) Indicatory origin—the zero point of the scale  $= {}^0I$   
(Eq. 20e, Ch. III)

(6) Indicatory range—the dispersion or limits  $= {}^sI$  or,  ${}^{a:z}I$   
(Eq. 20f, Ch. III)

b. Reliability—consistency on repetition under varying conditions. (See S. 48, Ch. X)

(1) Adequacy of sampling—dependent on the number of cases observed:

*Measured by:*

the standard error of sampling of the index

$= {}^s(I)$   
(Eq. 21a, Ch. III)

a measure of the amount of inconsistency from an infinite number of samples of this size

- (2) Observer error (or its complement of observer's reliability)

*Measured by:*

the significance ratios of the difference in means between 2 or more observers for the constant or one-way errors

$$= {}_{\sigma}({}^{p'-p''}I) \quad (\text{Eq. 21b, Ch. III})$$

the correlations between 2 or more observers for variable or two-way errors

$$= ({}^{p' \bullet p''}I) \quad (\text{Eq. 22a, Ch. III})$$

- (3) "Observand" error, i.e., error in the instrument-test, schedule card, questionnaire, interview procedure:

*Measured by:*

the significance ratio for constant errors,

$$= {}_{\sigma}(I_{/-//}) \quad (\text{Eq. 21c, Ch. III})$$

the correlation coefficient for variable errors

$$= I_{/\bullet//} \quad (\text{Eq. 22b, Ch. III})$$

- (4) "Observee" error,<sup>28</sup> i.e., error (or its complement, reliability) due to the responder, or informer.

*Measured by:*

the significance ratios and correlations between one responder on different dates (trials, re-observations) or between different informers in the plurel under observation (family, firm, club, etc.)

$$\left. \begin{array}{l} \\ \\ \\ \end{array} \right\} = {}_{\sigma}({}^{t'-t''}I) \text{ or } {}_{\sigma}({}^{p'-p''}I) \quad (\text{Eqs. 21d, e, Ch. III})$$

and

$$({}^{t' \bullet t''}I) \text{ or } ({}^{p' \bullet p''}I) \quad (\text{Eqs. 22c, d, Ch. III})$$

- c. Validity - correlation (including contingency) with criteria, i.e., with other indicators of the characteristic under observation. Percent of common elements or communality with one or more such criteria:

*Measured by:*

$$\left. \begin{array}{l} \text{the correlation coefficient of zero or-} \\ \text{der, by the multiple correlation co-} \\ \text{efficient and by the variance of the} \\ \text{communality} \end{array} \right\} = I_{\cdot\cdot\cdot}, I_{\cdot\cdot\cdot i}, h^2 \quad (\text{Eqs. 23a, b, c, Ch. III})$$

(See S. 1, 4, 8, 13, 14, 16, 17, Ch. VI)

Special case—Representativeness of Sampling,

*Measured by:*

$$\left. \begin{array}{l} \text{the correlation between category-} \\ \text{frequencies in the sample and in} \\ \text{the universe sampled (if known)} \end{array} \right\} = I_{i\cdot j} \quad (\text{Eq. 23d, Ch. III})$$

This outline of what calibration of indices means is not exhaustive, but is intended to make the student of Sociology realize the many inadequacies of most current societal indicants and the necessity of mastering the methodology of scale construction, field work, and statistics, if the observed data of sociology are to increase in precision. This improvement<sup>29</sup> in measurement is the fundamental task of any science. Theory is built upon it. Prediction and control depend far more on the accuracy of the observed data than on their statistical manipulation.

This classification of indicators on the basis of their precision is a classification based on the properties of the unit (as denoted chiefly by the class-interval script). But it is not solely a static classifying, it may also be a genetic one. Observations which were at the attribute level for cavemen may have reached the ordinal level by the Middle Ages and (aided perhaps by the introduction of the Arabic-Hindu system of decimal digits) may have reached the cardinal and calibrated stages only recently. In other words, observation of many phenomena may begin with noting qualities and ranks and become refined to calibrated indices.

While some phenomena occur in nature in a form readily expressed in cardinal units, such as head of cattle, most phenomena require the invention and agreement upon some objective standards before cardinal units emerge. Units are not inherent in the phenomenon; they are in the measuring instrument, or observer's mind. Accordingly, the hypothesis of degrees of precision of measurement, outlined above, is a working hypothesis for the scientist in that it challenges him to find out whether observations that are now at any qualitative or roughly quantified stage can be made more precise. The S-theory aids in this by specifying in its notation the operational stage of precision to which the several indices of the situation have been developed.<sup>30</sup>

### *C. Comparability of Units*

The process of developing calibrated indices is often a long and arduous one.<sup>31</sup> It need not go through all the seven levels sketched above, nor go through them in sequence. It begins with exhaustive qualitative analysis, proceeds with the invention of objectifying indicators, goes on to their combination with suitable weights, and ends with their verifying by means of many calibrated tests.<sup>32</sup> Then, even when cardinal units have been devised, there remains the problem of their vast diversity. There are thousands of different kinds of units in use for societal phenomena. To get common or at least comparable units there are various technics which may be summarized as (see Ref. 35, pp. 109 ff.):

1. Percentages—on observing the frequency of occurrence of all-or-none indicants, and calculating the mean and multiplying by 100, there results a roughly comparable unit called a percentage.
2. Ranks—by expressing quantities as ordinals (1st, 2nd, and 3rd, most, more, less, least; quartiles, deciles, percentiles) rough comparability is secured.
3. Money—to exchange goods and services they are evaluated and expressed in cardinal monetary units.
4. Index numbers—to compare changes in diverse phenomena each may be expressed as a percentage of itself on a common date or period taken as the base, or as a ratio to a common variable.
5. Standard deviations—diverse abilities may be compared by

expressing them in terms of probability, or of their frequency of occurrence in the same (or standard) population.

6. Components and elements—the technics of component or factor analysis in Psychology and the as yet undeveloped further resolution of these into minute numerous uncorrelated elements is a new technic. (See Ch. VI.)

The last four types of units (#3–6) are cardinal units but may, of course, vary greatly in excellence of calibration.

#### V. THE INDICATORY CASE SCRIPT, $\{I\}$

The case script, or point script, is the presuperdescript,  $\{I\}$ , denoting the number of specified aggregated cases, or specified points on the indicant scale. The prime in the case script, as usual, identifies a particular point.<sup>33</sup> \*

#### VI. THE GEOMETRY OF INDICATORS, $\bar{I}$

##### A. Vectors and Scripts

In expanding further the sketch of the geometrical interpretation of S-theory given in Chapter II, the terminology of vectorial algebra will be found convenient though not essential. Since few sociologists have had occasion to study vector theory, a simple outline of the elementary principles of line vectors is given in the footnote below where it may be ignored by the reader who is familiar with vectors.<sup>34</sup> †

In applying the principles of vector theory to the S-theory, note first, that a vector represents an index. A quality is represented geometrically by a point somewhere in societal n-space, and quantity is represented by a scalar, the length of the vector through that point and some other point of origin. Thus, an attribute is represented by a point which, together with the zero point representing the absence of that attribute, defines a unit vector,  $\bar{I}$ , while an indicant being an amount of some quality, or kind of thing, is represented by any vector,  $\bar{I}$ . Every vector is a product of a unit-vector and the number of scalar units in its length,  $\bar{I}^1 = \bar{I} I^1$ . Vector theory thus expresses a clean-cut mathematical relationship between the qualitative and the quantitative as a point determining a direction and as a length, respectively.

\* For Eqs. 24a–25b, Ch. III, see notes at end of the chapter.

† For Eqs. 26–31, Ch. III, see notes at end of the chapter.

Matrical situations,<sup>35</sup> such as a statistical tabulation, may be represented by a sheaf, or ray, of vectors, all emanating from a common zero origin. Thus, in S. 35, Ch. II, the three vectors,  $\bar{I}$ ,  $\bar{I}'$ ,  $\bar{I}''$ , in this figure might be the geometric representation of three of the indicants of home status in S. 6, Ch. II. Any array of any matrix may be called a row (or column) vector.

The indicatory class script states the number of indicatory vectors, just as, in general, the class script states the number of vectors in each of the other sectors. Thus, a single class script,  $I$ , would be represented by one vector, and a multiple script by the same number of vectors as the number of indices it denotes. Thus, for example, in S. 14, Ch. III, the formula  $(I)_i$ , where  $|_i = 2$ , denotes two indices and would be geometrically represented by two vectors. But these two resultant vectors (which may be correlated) can be resolved, as stated in the equations, into  $n$  uncorrelated components, i.e., into  $n$  orthogonal reference vectors. Each resultant vector can be expressed as a sum of a proportional part of each of the uncorrelated components called elements. The proportional part that each contributes is denoted by its loading or weighting coefficient,  $X$  (or  $Y$ ). Thus, in an orthogonal space of  $n$  dimensions (defined by the  $e$  elements), the  $S$ -situation defines two observed resultant vectors,  $x$  and  $y$ , in terms of the  $n$  hypothetical elemental vectors. Altogether in this situation there are  $n + i (= j + 2)$  vectors, which are the sum of the class scripts.<sup>36</sup>

The class-interval script states the scalar subdivisions (the units of length) of the vector representing the index. In the case of attributes each qualitative class is considered to be a unit-vector which is a vector of unit length. If there are subclasses these are considered the unit vectors, and the more inclusive class would be the resultant vector if the unit vectors were summed. If they are merely aggregated, the more inclusive class represents a sheaf of unit-vectors radiating out from a common origin. To consider, in a collection of qualitative entities, the value of each to be unity as is done by the zero exponent ( $I^0 = 1$ ) is, in effect, weighting the qualities as quantitatively equal. In the absence of any better weighting this is the most convenient and logical weighting to conventionalize as the standard. Whenever there is known to be, or may possibly be, another weighting, coeffi-

cients may be attached to the attributes in order to express this. Such coefficients convert the attribute into an indicant (or into a population, space, or period of *some kind*, if the coefficients are P, L, or T respectively).

The indicatory case script is geometrically represented by points on the vector. The multiple case script states the number of specified points; the singular case script (a primed letter) identifies one point. Since any point taken in conjunction with the origin defines an interval along the line determined by those two points, such a point *connotes* a distance as a *by-product* of *denoting* a point. Thus, in S. 15, Ch. III, the <sup>m</sup>I states that three mean points (centroids at A, B, and C,) are specified in addition to the other vectors.

In sum the vectorial formula for the indicatory sector is:

$$\begin{aligned} \text{the number of points} &= \overline{i} = \text{the number (less one) of} \\ &\quad \text{normals to each vector} \\ I &= \text{the scalar length, the num-} \\ &\quad \text{ber of I-units} \\ \text{the number of line-sects} &= i = \text{the number of vectors} \end{aligned}$$

(Eq. 32, Ch. III)

### B. Vectors and Operators

The vectorial meaning of the operational symbols is as follows:

The heavy dot, •, as usual in vector theory, denotes the scalar product of two vectors,

$$\bar{I}' \bullet \bar{I}'' = I, I'', \cos \theta \quad (\text{Eq. 28a, Ch. III}).$$

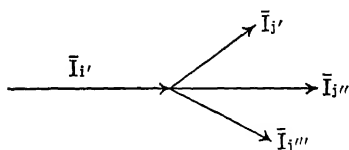
The scalar product of two unit-vectors is the cosine (or correlation coefficient) alone.

$$\bar{I}' \bullet \bar{I}'' = (1) \bullet (1) \cos \theta = \cos \theta = r$$

(Eq. 28b, Ch. III)

The colon, :, denotes subdivision of the vector in some way. Further exploration is required before its vectorial meaning can be precisely specified in all situations. Between class-interval scripts the colon denotes subdividing a vector into smaller line-sects, as in subdividing a series of years,  $\bar{t}\bar{I}$ , into years and months,

$t: u \bar{T}$ . Between class scripts the colon denotes that the vector becomes a ray of vectors. Thus  $\bar{I}_i' : j$ , represented as:



means that the ray of vectors,  $\bar{I}_j$ , are all in the general direction of  $I_i'$ , which may be their resultant (i.e., vectorial sum), or their centroid (i.e., a vectorial average). Between indices a variety of interpretations for the colon seems possible.<sup>37</sup> \*

### C. Dimensions Geometrically Defined

Turning from the geometric interpretation of the operational symbols, the chief remaining geometric interpretation of S-theory to note, is that of the class and exponent scripts. The class script denotes the number of vectors which is the apparent<sup>38</sup> number of dimensions of the s-space occupied by that S-situation. That is to say, if there are s indices, as shown by the sum of the class scripts in the four sectors, these s vectors apparently require a hyperspace of s dimensions.

$$|z_t + |z_i + (|^1)z_1 + |z_p = |z_s = \text{number of apparent vectorial dimensions in a situation} \\ \text{(Eq. 35a, Ch. III)}^{39}$$

If those dimensions are orthogonal,<sup>40</sup> i.e., mutually perpendicular, they are a set of s Cartesian co-ordinates.

The concept of "dimensions," meaning an s-space determined by the directions of the s vectors, may be referred to as "vectorial dimensions." Their number in each sector of a situation will be symbolized by  $|z'_s$ , and by  $|z_s$ , or simply s, for all sectors of the S.<sup>41</sup> †

This specifies the number of societal dimensions in one situation. But situations are arbitrarily delimited samples of societal data, and may range from situations of very minute and restricted scope, to situations of world-wide and permanent import. (E.g., compare the local, S. 1, Ch. III with the universal S. 6, Ch. III;

\* For Eqs. 33a-34, Ch. III, see notes at end of chapter.

† For Eqs. 36a-d, Ch. III, see notes at end of the chapter.

the particular S. 6, Ch. II with the general S. 5, Ch. II, etc.) Can our situations, our gross units of observation, be combined so as to yield some estimate of the dimensions of society as a whole? If the question is qualitatively put as, "What are the societal dimensions?" it can be answered moderately well, but if the question is quantitatively put as, "How many societal dimensions are there?" it cannot at present be answered, except by the roughest series of assumptions and approximations.

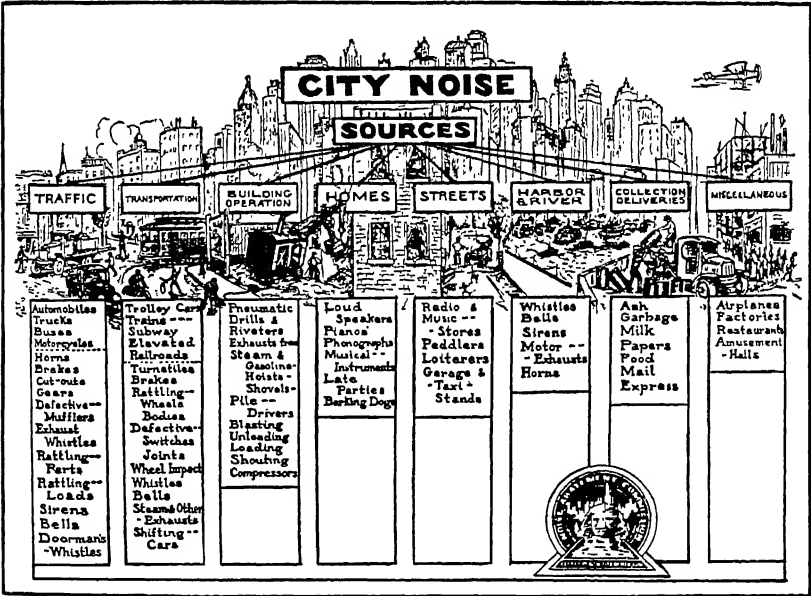
Dimensions, as here defined, are any distinguishable qualities whose amount, or at least whose presence or absence, is determinable. Dimensions pertaining to human society are conveniently classified into four classes for general systematic purposes (T, I, L, P), and into further subclasses for the more specialized purposes of the various social sciences. The number of *kinds* of dimensions may thus be conventionally agreed upon and so specified. The number of possible dimensions is an astronomical figure. It depends upon the following:

- a. every possible time series of dates or periods, grossly and finely subdivided, of any and every total duration from seconds to millions of centuries;
- b. every possible characteristic known to man. The classification of indicators by content above is a mere suggestion of the range of characteristics. The vocabulary of the largest dictionary is one rough estimate of characteristics that have been *named* for purposes of verbal communication;
- c. every possible spatial series of points, lengths, areas, volumes, single or many, small or large, everywhere and in every arrangement;
- d. all possible parties, including the two billion living people, plus the dead of the past, and every combination of them in pairs or larger plurels; and finally
- e. every possible variation, combination, and permutation of the above.

The salient fact is, not the enormous *possible* number, but the *actual* number which, relative to the possible total, seems an infinitesimal proportion—although it is still perhaps of the absolute order of  $10^{10}$  to  $10^{20}$  for a mere speculation. (See the estimation in Chapter VII for a closer estimate of a part of the field of societal dimensions.) But the total number of societal dimensions

is of little more importance than the equally rough estimate of the number of chemical molecules in the universe, when astronomers do not yet know the exact limits or size of that universe. We can study types of dimensions, their uniformities of functioning under classified and specified conditions, their number in particular defined situations, and much more, all of which cumulatively enables man to make better practical adjustments, to progressively understand, predict, and control phenomena better. This is the function of science. Beyond that, is still metaphysics which we leave to philosophy.

S. 1



Ref.: "New York's Noise," *Survey*, Vol. LXV, No. 4, Nov. 15, 1930, p. 217.

Descriptive formula:  $S_1 = I^0_{1;j}$       Quantic number = 0;0;0;0  
Legend:  
 $S_1$  = The situation is       $|_i$  = 8 classes by source  
a record of      and  
 $I^0$  = city noise       $|_j$  = 67 subclasses  
in

## S. 2



Ref.: Marshall, Leon C., *The Story of Human Progress*, The Macmillan Company, 1925, p. 314.

Descriptive formula:  $S_2 = I_1^1$

Quantic number = 0;0;0;0

Legend:

$S_2$  = The situation

: = to which correspond

= = is a record of

$|_1$  = 6 other institutional indicators

$I_1^1$  = a particular qualitative indicator, the school,

Comment on notation:

As usual the legend puts the formula into words for the benefit of those unfamiliar with the notation. The only essential information in the legend that is not in the formula is:

$|_1$  = the school

$|_2$  = 6 other institutions

## S. 3

## OBJECTIVES FOR AGRICULTURE

To Obtain:

- A. Recognition of the interdependence of rural and urban interests
- B. A favorable economic environment
- C. Efficient management and production methods
- D. Effective group action through organization
- E. A satisfactory social environment and standard of living
- F. Adequate rural educational opportunities

*Ref.: A Statement of Objectives for Agriculture, Iowa State College of Agriculture and Mechanical Arts, May, 1933, p. 5.*

*Descriptive formula:*  $S_3 = I_1^0 : ( , , , )$

*Quantic number* = 0;0;0;0

*Legend:*

$S_3$  = The situation

: = each qualified

= = is composed of

by

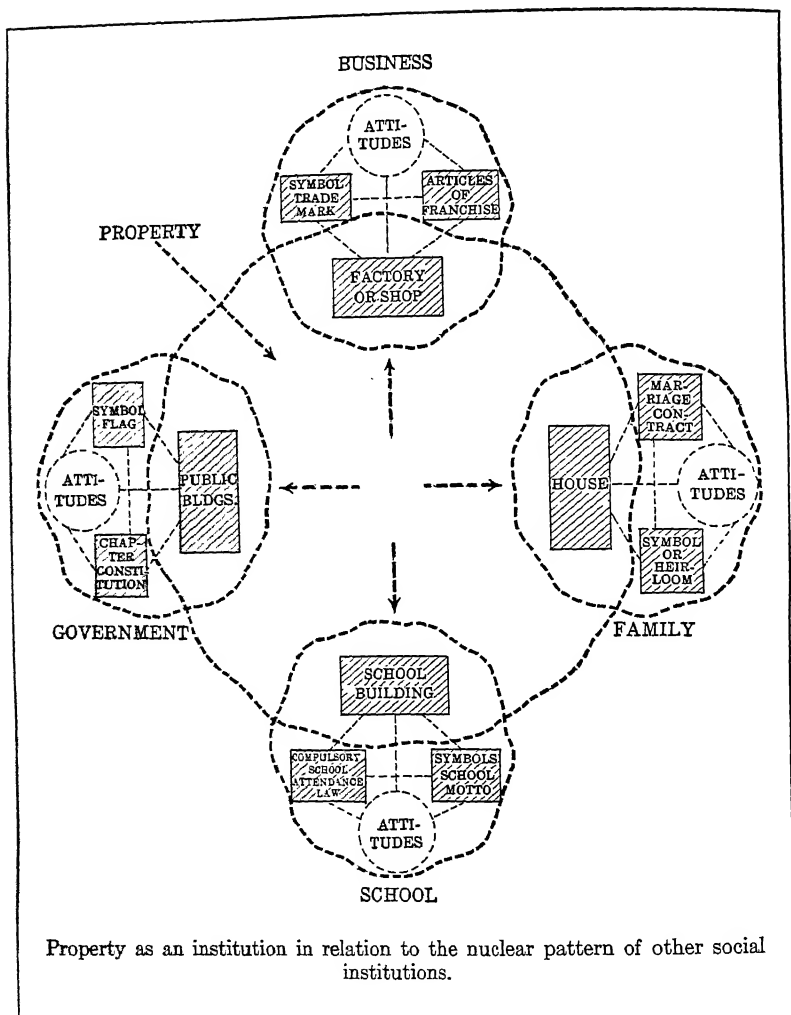
$I_1^0$  = 6 qualitative objectives of  
agriculture

$( , , , )$  = further attributes (adjectives)

*Comment on notation:*

Note that the qualifying adjectives and phrases are qualitative characteristics which are, therefore, attributes; and that multiplication of these qualitative indicators represents their meaning mathematically, as modifying, qualifying, or limiting the main attribute.

## S. 4



Ref.: Chapin, F. Stuart, *Contemporary American Institutions*, reprinted by permission of Harper and Brothers, 1935, p. 344.

Descriptive formula:  $S_4 = I_{1::j}^0$

Quantic number = 0;0;0;0

Legend:

$S_4$  = The situation

$::$  = cross-classified with

$=$  = is a record of

$|_j$  = 4 pattern-parts

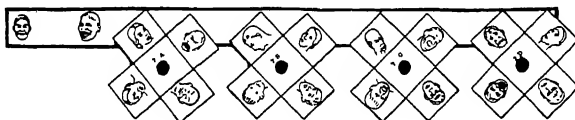
$I_1^0$  = 4 types of institutions

*Comment:*

Compare S. 19, Ch. II, which is essentially the same data in undiagrammed form. But they have been analyzed into different quantic formulae. S. 19, Ch. II writes the P explicitly, while here it is written as nul (P%), giving quantic numbers of 0;0;0;1 and 0;0;0;0 respectively. This is an example of the ambiguity still remaining in the Rules of Appendix, which permit unreliability of analysis such as this, and which produced the 1% to 3% of disagreement between independent analysts as reported in the experiment on the reliability of S-analyses in Chapter II. If the people, or plurel, implied in the terms "family," "school," "business," and "government," have been "canvassed or enumerated," the P must be explicit, but it is also possible to interpret the graph as presenting purely cultural phenomena ("attitudes," "symbols," etc.) with a very remote implication of the number of parties involved. Further research towards refining the Rules for S-analyses in order to reduce such unreliability as this is needed.

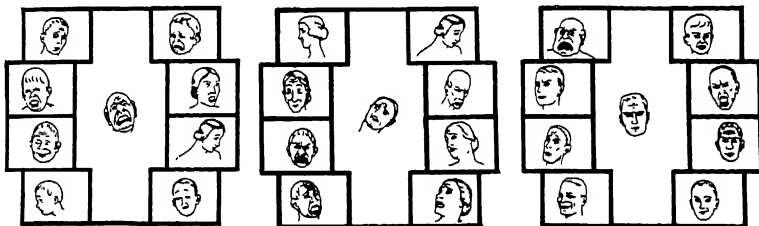
## S. 5

## TEXT 8 FORM A ("FACES")



First test row  
"Pleasure"

## TEST 4 FORM B (Revision, in paper and pencil form, of Test 8 above)



Third test problem

Ninth test problem

Sixteenth test problem

The task is a continuation of that in Plate X. In Form A, the four expressions which are most similar to the given two at the left are to be turned up; in Form B, the two expressions most similar to the central one are to be crossed out.

*Descriptive formula:*  $S_i = [iI^0 = ({}^iI^0)_i]_{j:k}$

*Quantic number* = 0;0;0;0

*Legend:*

$S_i$ = The situation	in each of
$[=]$ = presents an equation between	$ _i = 4$ † problems (i.e., on the rotating crosses) and all this is repeated
$iI^0 = 2^*$ cases of a qualitative characteristic	in each of
and	$ _i = 2$ forms of the test
$iI^0$ = one case of it	each having
mixed with	$ _k = 1$ and 3 items respectively
$iI^0 = 3$ cases of other characteristics	

*Comment on notation:*

1. Qualitative equivalences have long been used by psychologists in tests of "similarities" or "synonyms" which are found to measure criteria of intelligence. The attribute hypothesis of S-theory provides a notation which brings these qualitative equivalences under the usual mathematical rules for equations of quantities.

2. The situation is a complicated matrix equation of the fifth degree, as shown by the five multiple aggregative scripts,  $|_i$ ,  $|_j$ ,  $|_i$ ,  $|_j$ ,  $|_k$ .

The lowest degree is the aggregation of the 4 pictures (on a rotating cross in Form A, or on one side of a "box" in Form B) constituting possible answers  $|_i$ ; the second degree includes these and the central picture (in Form B or 2 left-most pictures in Form A) which set the problem  $|_j$ ; the third degree aggregates four such problems in Form A and two such problems in Form B (in the two sides of the "box" of 9 pictures)  $|_i$ ; the fourth degree aggregates three such test items in the 3 "boxes" of Form B,  $|_k$ ; and the fifth degree aggregates the 2 forms, A and B,  $|_j$ .

3. Since in S-theory qualities are geometrically represented by directions, the problem in this test of marking the pictures that are qualitatively most alike would be geometrically represented as marking the pictures whose unit vectors made the smallest angle with the unit vectors representing the pictures which set the problem. (Each picture is an attribute and every attribute is representable by a unit vector. The unit is the presence-to-absence of the attribute.)

4. The entire situation has 45 pictures, or 45 attributes, and therefore, 45 unit-vectors and 45 vectorial dimensions. The situation is operationally so defined (by assigning one point of score to each right answer, or picture, correctly indicated) as to summarize these 45 dimensions in the one resultant dimension defined by the score of this "similarities" test. This score then, is fulfilling the parsimony function of science in reducing phenomena of many dimensions to expression in fewer dimensions. S-theory merely measures the ratio of such reduction as being 45 to 1 in this situation.

\* 1 in Form B (Test 4)

† 2 in Form B (Test 4)

## S. 6

## MAJOR CATEGORIES OF PHENOMENA IN GENERAL

WHERE MANIFESTED	I	II	III	IV
	<i>The known universe manifests itself as various forms of SUBSTANCE,</i>	<i>which, under the promptings of some effective CAUSATION,</i>	<i>undergoes CHANGE,</i>	<i>in course of which emerge various PRODUCTS</i>
In the WORLD of NATURE	<p><i>Physical Matter</i></p> <p>↓</p> <p>Electrons Atoms Molecules Masses</p> <p>These appear</p> <p>1. Singly, as <i>units</i> of their particular realm</p> <p>or</p> <p>2. Plurally, as <i>structural combinations</i> of single units.</p>	<p><i>Physical Causation</i></p> <p>↓</p> <p>This involves</p> <p>3. <i>Physical Energy</i>, which activates substance, but which is subjected to</p> <p>4. <i>Physical Control</i>, by means of numerous factors which determine the direction and form of physical movement.</p>	<p><i>Physical Change</i></p> <p>↓</p> <p>This involves</p> <p>5. <i>Physical Action</i>, (<i>process</i>) or movement of substance, which is revealed in some alteration of</p> <p>6. <i>Physical Relationship</i>.</p>	<p><i>Physical Products</i></p> <p>↓</p> <p>7. These take the form of reorganized substances and new states of activity, which become determining factors in new cycles of change.</p>
In the WORLD of MAN	<p><i>Societary Composition</i></p> <p>This appears</p> <p>1. Singly, as <i>Single Human Beings</i></p> <p>or</p> <p>2. Plurally, as <i>Human Plurals</i> composed of combinations of single human beings.</p>	<p><i>Societary Causation</i></p> <p>This involves</p> <p>3. <i>Societary Energy</i>, or human motivations, whose forms of expression are subjected to</p> <p>4. <i>Societary Control</i>, or influences emanating from any portion of human society which determine the direction and form of human behavior.</p>	<p><i>Societary Change</i></p> <p>This involves</p> <p>5. <i>Societary Action</i> or processes of movement of human beings, which is revealed in some alteration of</p> <p>6. <i>Societary Relationship</i>.</p>	<p><i>Societary Products</i></p> <p>7. These take the form of reorganized and modified groups, together with their evolved <i>culture</i>, which become conditioning factors in new cycles of change.</p>

Descriptive formula:  $S_6 = I_{i::j:k}^0$

Quantic number = 0;0;0,0

Legend:

$S_6$  = The situation

records

:: = cross-classified with

$|_i$  = 4 classes ("I-IV")

$I^0$  = attributes of categories

and each cell is subdivided into

in

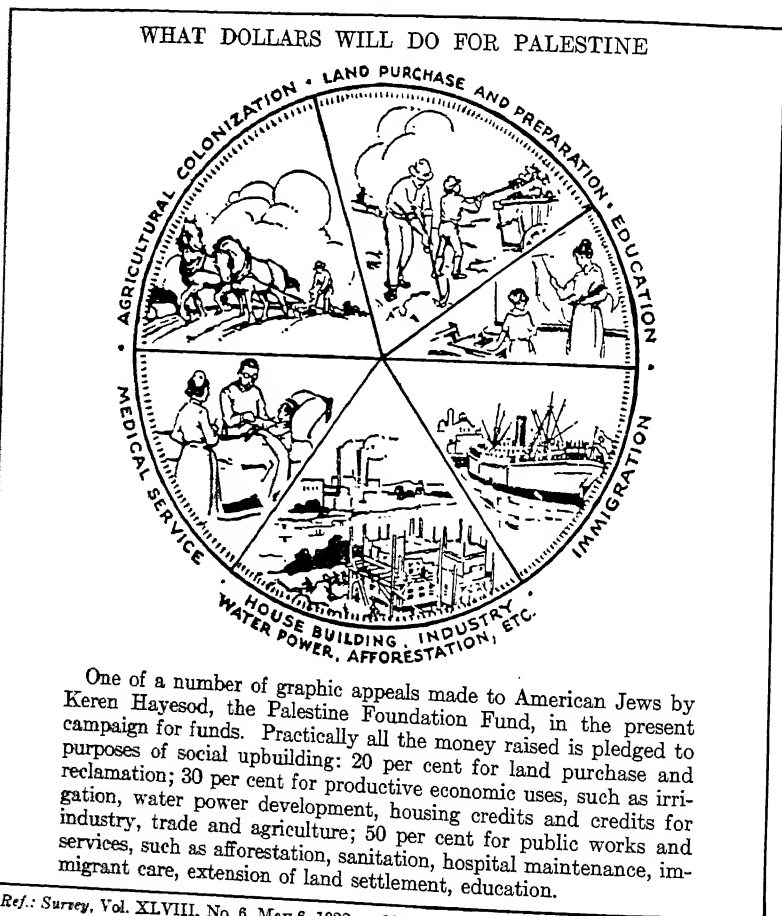
$|_k$  = 2 subclasses (as "1,2;3,4;5,6")

$|_i$  = 2 classes,  
i.e., "worlds"  $\left\{ \begin{array}{l} \text{"nature"} \\ \text{"man"} \end{array} \right.$

Comment:

The qualitative indicators, here, are words naming general ideas. Their tabulation in a matrix makes them expressible in an S-formula, better than if they were in a prose paragraph.

S. 7



*Descriptive formula:*  $S_7 = \%I_1$

*Quantic number* = 0;1;0;0

*Legend:*

$S_7$  = The situation

for

is composed of

$|_i$  = 6 classes of colonizing activities

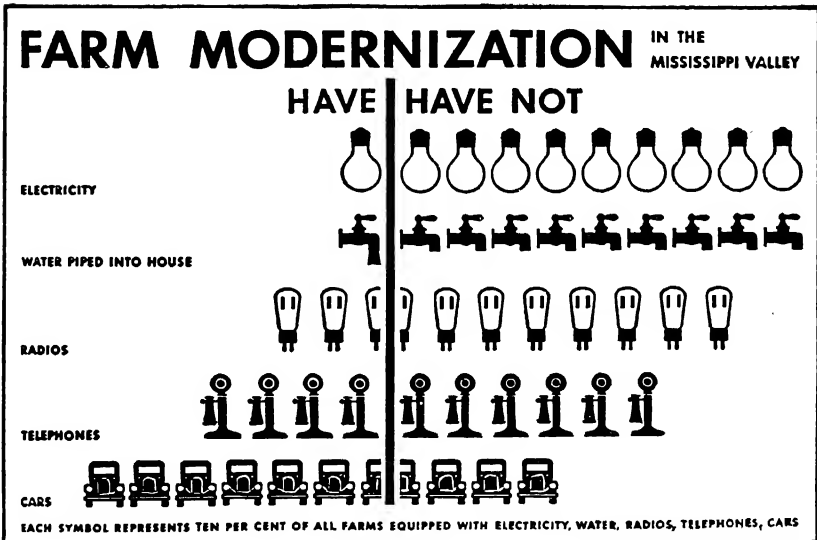
I = an indicant, dollars,

$\%|$  = expressed in percentages

*Comment on notation:*

This S-situation is a borderline one between a quantic number of 0;1;0;0 and 0;1;0;1. Is the reference to the population of Palestine sufficiently explicit to require writing the plural in the formula (i.e.,  $\underline{P}$ , :  $\%I_1$ ) and making its quantic formula,  $I^{+1}P^{+1}$ , denote a distribution? It was here ruled that, since no census or canvass of the Palestinian population is shown by the situation as presented, the population index is nul ( $P_0$ ), and therefore, the P is not written.

S. 8



Ref.: Smith, Russel, J., "The Sound Use of Land and Water," The Mississippi Valley Committee Report, *Survey Graphic*, Vol. XXIV, No. 2, Feb. 1935, p. 65.

*Descriptive formula:*  $S_8 = \%I_1 : j$

*Quantic number* = 0;1;0;0

*Legend:*

$S_8$  = The situation

$|_j$  = in 2 class- { "Have"  
intervals { "Have not"

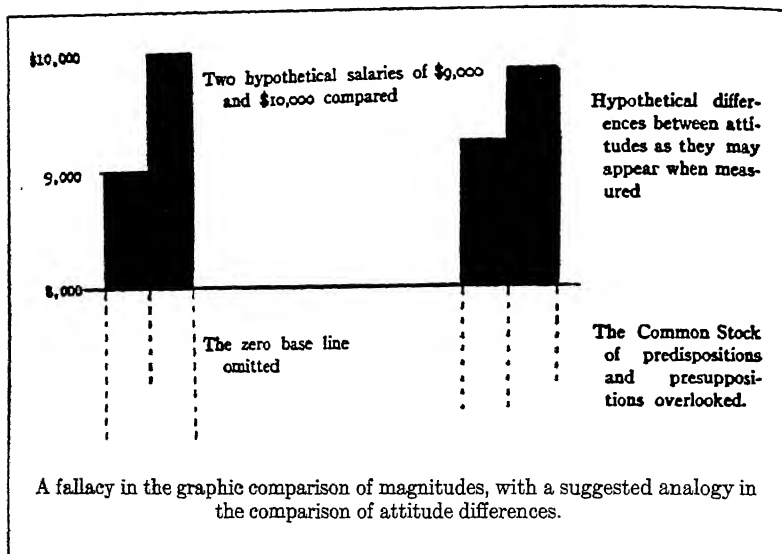
records

I = indicants of farm equipment

$\%|$  = each expressed as a percentage of their sum

$|_i$  = 5 in number

## S. 9



Ref.: Rice, Stuart A., *Quantitative Methods in Politics*, F. S. Crofts & Co., 1928, p. 120.

Descriptive formula:  $S_9 = \overset{0}{1}I$

Legend:

$S_9$  = The situation

records

$I_1$  = 2 indicants of stipend and attitude

Quantic number =  $0;1;0;0$

compared in

$|$  = 2 class-intervals

when

$\overset{0}{|}$  = the origin is indefinite

Comment on notation:

1. The digit 0 in the point script regularly denotes the origin or zero point of the scale of the component to which it is attached.

2. A borderline ruling is involved in the word "salary," which connotes an amount per period and hence would have a quantic number of  $9;1;0;0$ , putting this situation in the chapter on Change. But here the reference to the time component seems very remote. A salary status, an amount of money, is the more direct and simple implication of the comparison, and hence the formula is written with indicants only.

## S. 10

PROBABILITY OF THE RECURRENCE OF A DIFFERENCE SIMILAR  
TO THE OBSERVED ONE BETWEEN SUCCESSIVE QUINTILES AND  
EXTREME QUINTILES OF THE ENTIRE POPULATION IN ANOTHER  
SAMPLING OF HOMES

(Expressed by the critical ratio  $\frac{D}{\sigma \text{ diff.}}$ )

<i>Environmental Item</i>	<i>Quin- tiles 1 and 2</i>	<i>Quin- tiles 2 and 3</i>	<i>Quin- tiles 3 and 4</i>	<i>Quin- tiles 4 and 5</i>	<i>Quin- tiles 1 and 5</i>
1. Mother not employed.....	1.10	0.65	1.44	0.26	2.87
2. No relief before 1929.....	4.24	2.35	1.34	0.36	8.48
3. No relief after 1929.....	4.77	2.72	1.76	1.50	13.24
4. No juvenile court delinquency record...	2.76	3.02	0.99	1.78	6.66
5. No stores in block.....	1.56	3.01	2.44	0.28	7.33
6. Home not connected with store.....	0.33	0.68	1.85	0.08	1.50
7. No factories within 3 blocks.....	4.51	3.64	2.52	1.40	13.96
8. Parents not separated or divorced.....	1.17	2.28	0.83	0.00	2.57
9. Do not rent home.....	2.85	2.20	0.28	1.94	7.82
10. Central heating system.....	8.72	3.42	1.41	1.34	17.03
11. Telephone.....	8.28	4.63	3.44	0.92	29.66
12. Vacuum cleaner.....	10.75	4.45	1.86	0.55	21.32
13. Automobile.....	6.19	3.40	2.94	1.41	17.87
14. Second automobile.....	0.55	0.77	0.64	5.34	6.77
15. Radio.....	9.00	0.27	1.03	0.43	10.46
16. Piano.....	5.78	2.65	3.18	0.17	14.29
17. Second bathroom.....	2.48	0.77	3.21	3.40	6.45
18. Washing machine and mangle.....	0.54	0.99	3.40	2.08	7.41
19. Electric refrigerator.....	1.86	3.41	3.72	3.73	14.03
20. Folding camera.....	3.09	4.20	2.24	3.08	16.26
21. Moving picture camera.....	0.83	0.00	1.86	3.85	5.68
22. Playground equipment.....	3.02	2.76	2.54	4.49	15.76
23. Nursery or recreational room.....	1.86	2.58	1.64	4.46	10.47
24. Boat.....	2.48	1.64	0.17	2.22	6.15
25. Bicycle or tricycle.....	3.09	2.10	2.07	3.36	12.63
26. Typewriter.....	2.86	3.31	2.00	0.64	9.08
27. Fireplace.....	1.60	3.76	1.59	7.34	20.08
28. Ventilating fan for kitchen.....	0.10	0.61	1.22	1.44	3.57
29. Desk.....	3.93	1.57	2.17	0.71	8.78
30. Table lamps.....	0.69	0.36	0.28	2.10	4.38
31. Metropolitan newspaper other than local	1.60	0.40	1.17	1.00	4.18
32. Country or lake home.....	0.00	2.41	0.44	0.87	1.41
33. Vacations for family members.....	4.16	2.87	2.36	2.52	14.32
34. Either parent plays musical instrument.	1.46	2.77	3.70	0.50	9.24
35. House in good repair.....	6.11	1.90	1.09	1.67	11.28
36. Paid assistance in home.....	1.66	0.20	6.54	4.76	15.59
37. Two or more daily newspapers.....	3.92	2.13	1.84	1.55	9.77
38. Encyclopedia.....	5.07	4.44	1.70	1.82	16.91
39. Preventive dental treatment for child ..	4.10	2.01	3.83	1.87	13.97

## S. 10 (Continued)

Environmental Item	Quin- tiles 1 and 2	Quin- tiles 2 and 3	Quin- tiles 3 and 4	Quin- tiles 4 and 5	Quin- tiles 1 and 5
<i>Father's Membership in:</i>					
40. Professional or scientific society . . . . .	0.92	3.49	3.43	4.72	13.07
41. Civic club . . . . .	2.35	2.02	1.93	3.99	10.92
42. Trade union . . . . .	4.06	1.87	1.32	1.06	0.13
43. Parent-teachers association . . . . .	2.88	1.86	3.70	2.13	12.19
44. Study club . . . . .	0.48	0.43	3.23	2.68	6.41
45. Fraternal organization . . . . .	2.56	0.99	2.83	1.87	9.06
46. Social club . . . . .	3.87	2.58	2.51	4.19	16.40
47. University extension course . . . . .	0.92	2.97	0.53	2.86	5.40

Ref.: Leahy, Alice M., *The Measurement of Urban Home Environment*, Univ. of Minnesota Press, 1936, p. 32.

Descriptive formula:  $S_{10} = \sigma_{(p'-p'')^{\frac{MI}{I}}; i; j}$

Quantic number = 0;1;0;0

Legend:

$S_{10}$  = The situation

$p'-p''$  = the difference between 2  
quintiles

records for each of

in

$i$  = 47 environmental items

$MI$  = their mean scores

( $I$ ) = indices of probability

for each of

$\sigma$  = expressed in units of sigma  
of the difference

$j$  = 5 classes (pairs of quintiles)

the index being

## S. 11

## DATA ON STATEMENTS ACCEPTED

Statement	Mean Rated Position (Eleven- point scale)	Standard Deviation	Score Assigned
A. If I wanted to marry, I would marry one of them . . . . .	.25	.44	0
B. I would be willing to have as a guest for a meal . . . . .	2.6	1.46	25
C. I prefer to have merely as an acquaintance to whom one talks on meeting in the street . . . . .	4.9	1.26	50
D. I do not enjoy the companionship of these people . . . . .	7.4	1.00	75
E. I wish someone would kill all these individuals . . . . .	9.6	.47	100
F. I know nothing about this group; I cannot express an attitude . . . . .	*	*	*

\* Omitted from all calculations.

Ref.: Dodd, Stuart C., "A Social Distance Test in the Near East," *Amer. Jour. Soc.*, Vol. XLI, No. 2, Sept., 1935, p. 195.

*Descriptive formula:*  $S_{11} = {}_{i,j,k,l}I$

*Quantic number* = 0;1;0;0

*Legend:*

$S_{11}$  = The situation

and

records

${}_1|$  = a special class of attitude "I know nothing—"

$I$  = an attitude indicant

${}_{i,j,k}|$  = in 3 sets of units with 5 class-intervals each

*Comment:*

This situation records the five statements of attitude which were calibrated in constructing a social distance test. The "mean rated position" determines the intervals between statements converting the statements from ordinal into cardinal units.

## S. 12

IV. SANITATION ACTIVITIES						
A. GENERAL SANITATION (Total Points 100)						
45. Routine Inspection Service	Value of Item	Value Assigned				Total
		HD	OOA	PP	VA	
Number of sanitary inspections and re-inspections . . . . .	10					
Standard: 3000 inspections per 100,000 population						
Inspections—At least 3000 . . . 10 points						
2500 . . . 8						
2000 . . . 6						
1500 . . . 4						
1000 . . . 2						
Less than 1000 . . . 0						
Inspections . . . . . ÷ No. of hundred-thousands of population . . . . . = . . . . . inspections per 100,000 population						
46. Water (35)						
a. Local water meeting U.S. Treasury Department Standards for drinking water on common carriers and generally available (U.S. Treasury Department Standards adopted June 20, 1925) . . . . .	30					
Standard: 100 percent distribution						
Percent—At least 100 . . . . . 30 points						
70 . . . . . 0						
NOTE: The score on water supply is based on quality. Water which meets the U.S. Treasury Department Standards of purity for water on common carriers in Interstate Commerce, or which is certified as being equivalent to U.S. Treasury Department shall be given full credit. Example of scoring:						
City water meeting U.S. Treasury Department Standards and 100 percent distribution—score 30 points						
For each point under 100 percent deduct 1 point from score. A city with 80 percent distribution:						
100 - 80 = 20						
30 - 20 = 10 points—score allotted						
b. All cross-connections between city supply and other water supplies eliminated and all dual supplies inspected semi-annually . . . . .	5					
(No credit allowed unless inspected at least semi-annually)						

Ref.: Appraisal Form for Rural Health Work, American Public Health Association, 1934, pp. 4 and 41.

*Descriptive formula:*  $S_{12} = {}_{i,k}I_{i:j:k}$

*Quantic number* = 0;1;0;0

*Legend:*

$S_{12}$  = The situation

VA'') expressed in 2 kinds of units

records

$I_i$  = 2 types of sanitation activities  
("45" and "46")

$|_i$  = performance percentages and

$|_k$  = score-values

and

and

$|_i$  = 3 subtypes

$|_i$  = maximum or standard points  
are also noted

in each of

$|_k$  = 4 agency categories ("HD . . .

*Comment:*

This situation is a fragment of a scale which represents one of the most scientific and highly elaborated attempts yet made to reduce a culture complex ("public health activities") to quantitative summarization.

The specifications of this appraisal form are "operational definitions" of the I-units, as they direct the operator how to compute it from observed data.

## S. 13

## Armament or Improved Farming?



FARM TRACTOR

A tractor and an automobile for every one of the 6,500,000 farms in the United States could be bought with the money the great war cost the United States for a half year, and there would be left

600 million dollars for good roads.

Ref.: *Facts on Disarmament Exhibit*, Disarmament Education Committee, Washington, D.C., Card No. 9.

Descriptive formula:  $S_{13} = u \underline{I}_{\Sigma v = w}$

Quantic number = 0;1;0;0

Legend:

$S_{13}$  = The situation

in

records

$u$  = units of 1 per farm

$\underline{I}$  = the unstated total of money

= = equated to

needed to buy

$I_w$  = the cost of half a year of the great war

$\Sigma$  = 3 kinds of values

{ tractors  
autos  
good roads

*Comment:*

In addition to the equality between dollars, there are implied units of ethical evaluation in terms of which the equation might become an inequality of "greater than" or "less than," depending on the militaristic or pacifistic standards of the evaluator.

*Comment on notation:*

A borderline decision is involved, as other analysts might write the cost of the war for half a year as velocity,  $IT^{-1}$ , yielding a quantic number of 9;1;0;0 and shifting this situation to the chapter on Change. As the reference to the time rate of expenditure seems unimportant in this situation where the point is a timeless comparison of values, the time index was not written explicitly.

## S. 14

The precise form of distribution of the  $e$ 's, however, does not need to be specified. The features of the  $e$  hypothesis are:

(1) *Independent elements.* The observed normally distributed variables,  $x$ ,  $y$ ,  $z$ , etc., can be expressed as functions \* of  $n$  small independent elements, thus:

$$\begin{aligned} x &= X_1e_1 + X_2e_2 + \dots + X_ne_n \dots\dots\dots(6), \\ y &= Y_1e_1 + Y_2e_2 + \dots + Y_ne_n, \text{ etc.} \end{aligned}$$

(2) *One-or-none weights.* The coefficients,  $X$ ,  $Y$ ,  $Z$ , etc., are either unity or zero.

(3) *Unequal probabilities.* The  $m$  tests,  $x$ ,  $y$ ,  $z$ , etc., are determined by different proportions,  $p_x$ ,  $p_y$ ,  $p_z$ , etc., of the  $n$  elements; i.e., every coefficient such as  $X_i$  has a probability of  $p_x$  of being 1 and a probability of  $1 - p_x$  of being zero.

The first assumption is the conventional one of statistical analysis. The second is chiefly for simplicity. It seems likely that fractional, or unequal, weights could be assumed just as well though the mathematical treatment would be much more difficult. All the unit weights are assumed positive as negatively weighted elements are not needed for our present hypothesis. The inequality of the proportions,  $p$ , has been found necessary to explain hierarchy, or steep equiproportion, as explained in Section VI (2).

'Sampling' may refer to dealing with finite numbers of either the  $n$   $e$ 's, the  $m$  tests, or the  $N$  individuals. Ordinarily 'sampling' refers to the  $N$  individuals and such sampling error is measured by the probable error formulae. The essence of Thomson's Sampling Theory of Ability, however, deals with considering each test as depending upon a sample, or proportion, of the  $n$   $e$ 's.

\* This function is taken as an additive one on the basis of Taylor's theorem. The coefficients,  $X$ ,  $Y$ ,  $Z$ , etc., are the first order differentials and higher orders are neglected. Gauss used this assumption in deriving the normal distribution curve, Bravais used it in deriving the normal correlation surface, and it is the basis of probable error formulae. So the assumption may be considered a well-tried and accepted one in many sorts of biometric problems. Criticism is invited as to the conditions which may exist bearing upon its applicability here.

It should be noted that, if the coefficients are expressed in units such as dividing each by the sum of their squares (which with one-or-none coefficients is simply  $n$ ) they become correlation coefficients of that  $e$  with its dependent variable,  $x$ . They also are then the  $n$ th order regression coefficients of  $x$  on the  $e$ 's, for the regression and correlation coefficients are identical when the variables, such as the  $e$ 's, are uncorrelated.

Descriptive formula:  $S_{14} = [I = \sum_1^j JK]$

Quantic number = 0;1;0;0

Legend:

$S_{14}$  = The situation  
records

[=] = an equation between  
each of

$I_1$  = 2 variables (x,y)  
and

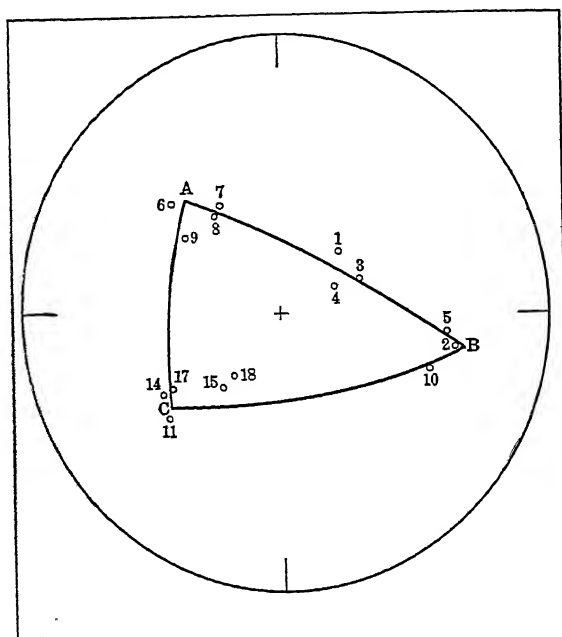
$\sum_1^j$  = the sum of indices, j in  
number (n)

each a product of

J = a weight or coefficient (X)  
and

K = an elementary variable (e)

S. 15



Ref.: Thurstone, L. L., *The Vectors of Mind*, Univ. of Chicago Press, August 1935, p. 164.

Descriptive formula:  $S_{15} = \bar{m}_1 \bar{I}_1 : j : k$

Quantic number = 0;1;0;0

Legend:

$S_{15}$  = The situation  
is composed of

$\bar{I}_1$  = 3 subcomponents (A, B, C)  
representing

$|_1$  = 4 clusters

$|_k$  = 15 vectors of indicants  
and

$\bar{m}_1$  = the mean of the ray of vectors  
i.e., the centroid, is also indicated

of

*Comment:*

The uncaptioned situation is a diagram of the surface of a unit-sphere showing the termini of unit-vectors roughly grouped into 4 clusters. The situation is a sociological one only when, as here, the vectors represent indices of human characteristics.

## VII. NOTES

1. For reference purposes, the equations are numbered consecutively within each chapter with the abbreviation "Eq." Most of the equations are simply definitional, defining a concept, or quantity, or function of them, in symbols.

2. It is possible to use three, five, or other sets of sectors, provided they are so defined as to include the entire societal field without overlap. Thus, for many studies, space can be neglected. On the other hand, in economic studies, for example, monetary indices are so basic as to justify denoting them by a separate symbol and using five sectors. Such a monetary index is still a subclass of characteristic, as here defined, however, and is simply labeled more explicitly for convenience without affecting the logic of our analysis into four basic types of indices.

Indices of time and space have always been recognized as basic in the history of philosophy and science. Special sciences have selected other additional indices as basic and have thereby marked off their field, as Physics did in adding mass to space and time. Economics analyzes production into the factors of land, labor, capital (and sometimes management), all operating in time, of course. These factors in S-terms are those of time, space, population and the *two* characteristics of capital and management which are peculiar to that field.

Statistical textbooks trace statistical series to variations in time, space, quality, and quantity. Our attribute hypothesis combines quality and quantity as subvarieties of characteristics. The population sector is added to *delimit* the sociological field, for the only major index, common to all the social sciences and shared by none of the subhuman fields, is the human population. The concepts of S-theory may readily be extended to all biology and other non-human fields, if desired, by defining suitable "characteristics" to be specially symbolized. But the field of this volume is human society.

3. An alternative sequence of topics may be preferred by some teachers. Such a one is to start, as some sociologists do, with the topic of population in its spatial, temporal, and qualitative distribution. (This is the  $P^{+1}$  array in the  $L^2$ ,  $T^{1,2}$ , and  $I^0$  cells of the quantic solid S. 33, Ch. II.) This leads on logically to quantitative characteristics of populations in Distributions, and then to their Correlations and Interrelations. Either the temporal or the space dimension could then be followed, cumulating first one and then the other, until the complete quantic formula for analyzing any and all situations has been developed. Still another and more mathematical sequence would be:

1. Zero order (nullary) indices,  $I^0$ ,  $T^0$ ,  $L^0$ ,  $P^0$ , especially qualitative indicators  $I^0$ ; then

2. First order (primary) indices,  $I^1$ ,  $P^1$ ,  $T^1$ ,  $L^1$ , quantitative indicators, plurals, and change, singly and in combinations; and then the more profound analysis of

3. Second order (secondary) indices,  $I^2$ , correlations of characteristics,  $P^2$ , interrelations of people, and  $T^{-2}$  accelerations of change.

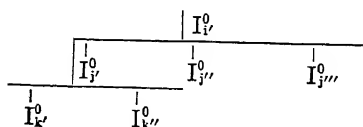
This sequence follows a diagonal route through the quantic solid, starting from the simplest situations in the bottom front leftmost corner and ending with the most complex in the top hind rightmost corner in S. 33, Ch. II.

4. The symbol for the indicatory exponent is  $|^i$ . Since, however, the letter  $e$  is widely standardized in mathematical usage to denote an exponent, it may be used interchangeably with any post-superscript (i.e., exponent) of S-theory.

5. The exact rules for adding, subtracting, multiplying, and dividing nullary indicators,  $I^0$ , are still being worked out by a process of induction from the recorded societal situations which have been reduced to S-formulae. As such situations cumulate, alternative rules as hypotheses are being tried out to see which fit the data best, and should therefore, become standard operational procedure in combining nullary indicators. This research is as yet only begun.

One tentative set of rules is as follows:

1. To add two attributes,  $I_j^0, I_{j''}^0$ , find the attribute of which the two are subclasses, e.g., dogs + horses = animals; red + green = colors; running + walking = locomotion; truth + beauty + goodness = values. To find the sum, go upwards in a hierarchical classification until a class is reached which subsumes the classes to be added.



Thus:

$$I_{j+j''+j'''}^0 = I_i^0 \quad (\text{Eq. 4a, Ch. III})$$

$$I_{k+k''}^0 = I_j^0 \quad (\text{Eq. 4b, Ch. III})$$

$$\therefore I_{k+k''+j''+j'''}^0 = I_i^0 \quad (\text{Eq. 4c, Ch. III})$$

Names have been invented for only those few sums of attributes (out of the possible billions) which are frequently used. Thus, we rarely need a name for such an arbitrary and unusual class as that defined by "a hammer + intelligence + fairies," although a word meaning the class of these three concepts would be invented were it frequently needed.

2. To subtract one attribute from another is to deduct the first from their sum, e.g., animals - dogs = non-dog animals; values - truth = values-other-than-truth. If the superior class is exactly the sum of the subclasses without other subclasses, then the subtraction of one subclass leaves the others, e.g., male humanity + female humanity = humanity  $\therefore$  humanity - male humanity = female humanity. Subtraction implies a previous sum. Rules for handling a negative attribute could doubtless be devised in some consistent way. In symbols, assuming Eq. 4a:

$$I_i^0 - j' = I_{j''+j'''}^0 \quad (\text{Eq. 5, Ch. III})$$

3. To multiply two attributes is to find the attribute defined by both the factor attributes simultaneously:

$$I^0 \times I^0_{/} = I^0_{+,,} \text{ or } I_{,,,} \quad (\text{Eq. 6, Ch. III})$$

pocket  $\times$  timepiece = watch  
 dead  $\times$  creature = corpse  
 woman  $\times$  mate = wife  
 swift  $\times$  motion = swift motion

Adjectives modifying nouns or other adjectives, and adverbs modifying verbs, are common examples of these qualitative products. Another suggestive side-light (discussed further in Ch. VI) is that in compounding the probability of independent qualitative characteristics, the probability of *either* of two characteristics is the sum of their separate probabilities,  $p + q$ , while the probability of *both* characteristics occurring jointly is their product,  $pq$ . If negroes are 20% of a population and women are 50% of that whole population, the probability of meeting either a man or a woman is  $50\% + 50\% = 100\%$ , while the probability of meeting a negress is  $20\% \times 50\% = 10\%$ .

4. To divide one attribute into another seems to specify "that-which-relates" those two attributes. Thus in the pictorial analogies test, S. 14, Ch. II, there is an equation between two ratios of qualities (i.e., attributes) which are verbalizable as:

$$\frac{\text{head of a man}}{\text{a whole man}} = \frac{\text{head of a dog}}{\text{a whole dog}} = \text{the head-body}$$

relation, a special case of the part-whole relation.

$$I^0_{/i'} = I^0_{i'/i'} \quad (\text{Eq. 7, Ch. III})$$

Every primary index in the 300 graphs (S-situations) in this volume is a product of an attribute and a "pure" duration, amount, length, or population, *of some kind*. Every primary index as formally defined by Eq. 9, Ch. III, Eq. 4, Ch. IV, Eq. 5, Ch. VIII, and Eq. 8, Ch. IX, is a product of a quality and a quantity. Every hierarchical descript in any of the 300 situations (i.e., of the form  $|i:j:k|, |p:q:r|, |l:m, t:u|$ , etc.) seems to imply addition or subtraction of qualities as one moves to the left or to the right respectively in the series of descripts. A "type" may prove analyzable as a qualitative average.

The few instances of division by attributes,  $I^0$ , will be commented on as they occur in the S-situations. (Until we discover more instances we cannot suggest, with any assurance, a generalization on qualitative division.)

A problem for further research is to relate and integrate, if possible, these symbols and rules with those of the symbolic logicians. As one hypothesis, test the apparent correspondence of the symbols which are the basis of Whitehead and Russell's *Principia Mathematica* (Ref. 81, Vol. 1, p. 6). (The scripts  $i + j$  below, are simplest when in the singular,  $i' + j'$ .)

1. The contradictory function,  $\sim$  may be relatable subtracting attributes,  
 $(\sim i \text{ means not } -i)$  to the S-theory  $I^0_1 \sim i? = -I^0_1$   
 operation of (Eq. 8a, Ch. III)

2. The logical sum, or disjunctive function,  $\vee$   
( $i \vee j$  means  $i$  or  $j$ )      may be relatable to the S-theory operation of adding attributes,  $I_i^0 + I_j^0$  or  $(I_{i+j}^0)$   
 $i \vee j? = I_{i+j}^0$   
(Eq. 8b, Ch. III)
3. The logical product, or conjunctive function,  $\cdot$   
( $i \cdot j$  means  $i$  and  $j$ )      “ multiplying attributes,  $I_i^0 I_j^0$  or  $(I_{ij}^0)$   
 $i \cdot j? = I_{ij}^0$   
(Eq. 8c, Ch. III)
4. The implicative function,  $\supset$   
( $i \supset j$  means  $i$  implies  $j$ )      “ aggregating attributes  $I_i^0 : I_j^0$  or  $(I_{i:j}^0)$   
 $i \supset j? = I_{i:j}^0$   
(Eq. 8d, Ch. III)
5. Equivalence,  $=$   
( $i = j$  means  $(i \cdot j) = (j \cdot i)$ )      “ cross-classifying  $I_i^0 :: I_j^0$  or  $(I_{i::j}^0)$   
 $(i = j)? = I_{i::j}^0$   
(Eq. 8e, Ch. III)

If our symbols should prove to correspond exactly to Whitehead and Russell's, so that the hypotheses stated in Eq. 8, Ch. III, are proven, then the sociological formulae of this volume might be transposed into their notation. If imperfect correspondence is revealed (after applying both sets of symbols to a large common sample of referents) it may prove desirable in the interests of the unity of science to modify the definitions of some symbols in either set in order to obtain a single set of symbols which can deal both with the logical propositions in the *Principia Mathematica* and also with the societal statistics in the *Dimensions of Society*.

6. An “implicit” index is to be distinguished from an “indefinite” index and from a “nul” index. An implicit index is a nullary one with non-zero descripts, and is not written explicitly in the full formula, S, but diphthonged as a product with another index. A nul index is one with a zero exponent and zero descripts. It, therefore, affects the situation in no way, not even qualitatively, as does the implicit index. The nul index is not written in the “full” or “descriptal” formula, S; it is written as zero in the quantic number.

The “indefinite” index, symbolized by underlining, denotes one that is asserted to exist, but its amount is not stated in the case of a primary index,  $\underline{I}^{+1}$ , or a multiple script,  $\mathbb{I}_i^1$ . In the case of a nullary index,  $I^0$ , and of a singular script  $\mathbb{I}_i^1$ , the underlining means that the identity is not stated.

	Exponent	Descripts	In S-formula
Explicit index	e	+	(I)
Indefinite index	e	+	( $\underline{I}$ )
Implicit index	0	+	not written
Nul index	0	0	not written

where e denotes any positive, negative, or zero exponent.

7. The symbolic formula for an index is:

$$(I) = \sum_1^z \mathbb{I}^1(I')^e = \text{an index} \quad (\text{Eq. 10, Ch. III})$$

When  $e = 0$ , (I) becomes a sum of  $z$  indicators

When  $z = 1$ , (I) becomes simply a product of the indicator and another homosectoral index, i.e., T, I, L, or P

When  $i = 0$ , the indicator becomes an attribute, qualitatively modifying the homosectoral index, (I').

8. The index number is symbolized in full as  $I^{+1}I^{-1}$ , or more compactly as  ${}_eI$ , where the class-interval script,  ${}_e$ , denotes that the indicant is expressed in percentage or relative units:

$$I/I = {}_eI = \text{an "index number"} \quad (\text{Eq. 11, Ch. III})$$

In the quantic formula which ignores the descripts, this is condensed to  $I^{+1}$ , as the divisor,  $I$ , is of the same qualitative kind of unit and merely shifts the size of the units.

9. If indicants of different classes are to be identified the convention is to use successive letters of the alphabet—thus,  $I_i$ ,  $I_j$ ,  $I_k$ , etc.

10. Some convenient refinements of notation may be noted by the advanced student of S-theory:

	<i>Indicators</i> ( <i>indicatory indices</i> )	<i>Sectors</i> ( <i>homosectoral indices</i> )	<i>Situations</i> ( <i>heterosectoral indices</i> )
1.	$ _i$ = the aggregation of indicators, $i$ in number	$ '_s$ = the aggregation, or pattern, of class scripts in any one sector, i.e., $ _{i;j;k}$ or $ _{p;q}$ OR $ _{l;m;n}$	$ _s$ = the aggregation, or pattern, of class scripts in all sectors, i.e., $( _t;  _{i;j;k};  _{p;q};  _{l;m;n}) =  _s$
2.	$ \Sigma_i$ = the number of indicators, the number of indicator vectors or dimensions	$ \Sigma'_s$ = the number of classes in one sector, the number of vectorial dimensions, the order of the matrix, i.e., $\Sigma_i + \Sigma_j + \Sigma_k = \Sigma_s$	$ \Sigma_s$ = the number of classes in all sectors, i.e., $\Sigma_t + \Sigma_i + \Sigma_j + \Sigma_k + \Sigma_p + \Sigma_q + \Sigma_l + \Sigma_m + \Sigma_n = \Sigma_s$ the total number of vectorial dimensions, the total order of the matrix
3.		$ \Sigma'_s$ = the number of class scripts in one sector, $\therefore$ the number of levels in a classification. Thus, in $ _{i;j;k}$ $ \Sigma_s = 3$ $ _{p;q}$ $ \Sigma_s = 2$ , etc.	$ \Sigma_s$ = the number of class scripts in all sectors (= 9 in the illustration just above, the degree of the matrix)
4.	$ _{i'}$ = the $i$ th particular or identified indicator	$ '_{s'}$ = a particular class script in a particular sector, such as $ _j$ or $ _q$ from the above	$ _{s'}$ = a particular class script for an index from any combination of sectors

11. In analyzing a situation as to its class script a refined notational point should be observed. Situations may have two kinds of classifications—those in common units and those in diverse units. (a) Thus, a situation may have several indicators, denoting as many kinds of things, and all expressed in one kind of unit. An example is S. 7, Ch. III, where the dollar units are used for six purposes in colonizing Palestine. (b) A situation may be composed of a variety of indicators which are not expressed in a common unit. An example is S. 31, Ch. II, where the situation of the Russian 5-Year Plan has two major indicatory indicants—freight and capital—which are expressed in tons and ruble units respectively.

The notation differentiating these kinds of classifications is:

$|_1$  = the list of different indicators,  $|_1$  in number; a measure of “dissimilarity” when summed as in  $|_{\Sigma}$ . (Eq. 14a, Ch. III)

$I_1$  = different indicants,  $|_1$  in number, expressed in the common units,  $I$ . (Eq. 14b, Ch. III)

$..I_1$  = different indicants each with its corresponding units, i.e., sets of class-intervals (Eq. 14c, Ch. III)

The quantic formula,  $I^{-1}$ , is the same for both, as it is the highest exponent in each of the two types.

12. A suggested research project would be to determine experimentally the reliability of this (or any other) classification (as was done for the quantic classification). By collecting mechanically samples of situations, having them classified independently by more than one person, and then determining the percentage of discrepancies, the ambiguous spots in a classification can be scientifically determined and progressively eliminated by refining the categories or their definitions.

13. Topographical indicators (points, distances, altitudes, areas, slopes, volumes) are the length component,  $L$ , with varying exponent,  $(L^1)$ .

14. When time as a duration or a velocity is explicitly included, the indicator becomes an index. But amounts of change can be measured without stating dates or time intervals.

15. Phrases such as this, “relatively untrained,” are examples of the unsatisfactoriness of this classification. What is the boundary between “relatively untrained” and the implied category “relatively trained”? All motor co-ordinations have been practiced, however incidentally, since birth, and therefore, are trained to that extent. To determine what amount of incidental training is below the average amount is impractical. An alternative is to measure individuals who have reached their final plateaus of learning with maximal practice as an indicator of native ability. This, however, is influenced by individual differences in motivation and other factors.

The purpose of the list presented here is, however, not so much to offer a classification satisfactory to the psychologist as to suggest to the student of Sociology the range of data from all the sciences which may have societal significance.

16. The number of persons or of plurels is the population dimension,  $P$ . The

qualitative and quantitative characteristics identifying or describing persons or plurals are the only indicators. The distinction is fully expounded in Chapter IV on Plurals.

17. This value class of indicators is the most important for Sociology. The evaluations of the group, conscious and unconscious, constitute the motivation of all societal action and the cause of most societal forces and of all societal control. So important is the concept of "value" that it will recur frequently throughout this volume in the form of the word "desideratum," which is more definite and has fewer ambiguous connotations. The determination of what indicators are indicators of desiderata to specified people at specified times is dealt with in Chapters V and X.

18. The topics starred are Small's "six interests."

19. For alternative classifications see Ref. 25, especially Chapters X and XV.

A suggestive alternative which, while less detailed for an industrialized culture, is common to all cultures, is given by a cultural anthropologist (Ref. 81, p. 74) as follows:

- |  |                                      |
|--|--------------------------------------|
| 1. Speech                                      | b. Treatment of the sick             |
| Languages, writing systems, etc.               | c. Treatment of the dead             |
| 2. Material Traits                             | 6. Family and Social Systems         |
| a. Food habits                                 | a. The forms of marriage             |
| b. Shelter                                     | b. Methods of reckoning relationship |
| c. Transportation and travel                   | c. Inheritance                       |
| d. Dress                                       | d. Social control                    |
| e. Utensils, tools, etc.                       | e. Sports and games                  |
| f. Weapons                                     | 7. Property                          |
| g. Occupations and industries                  | a. Real and personal                 |
| 3. Art—carving, painting, drawing, music, etc. | b. Standards of value and exchange   |
| 4. Mythology and Scientific Knowledge          | c. Trade                             |
| 5. Religious Practices                         | 8. Government                        |
| a. Ritualistic forms                           | a. Political forms                   |
|  | b. Judicial and legal procedures     |
20.  $\underline{T}^0; \underline{I}^0; \underline{L}^0; \underline{P}^0$  = quantic formula for attributes (Eq. 15a, Ch. III)  
 $0; 0; 0; 0$  = quantic number for attributes (Eq. 15b, Ch. III)  
 $\underline{T}^0; \underline{I}^1; \underline{L}^0; \underline{P}^0$  = quantic formula for indicants (Eq. 15c, Ch. III)  
 $0; 1; 0; 0$  = quantic number for indicants (Eq. 15d, Ch. III)

21. As to the best number of class-intervals to use, one formula (which tends towards making the frequencies proportional to the binomial coefficients and hence nearest to a normal curve) is (Ref. 71),

$$\Sigma_i | = 1 + 3.3219 \log_{10} P$$

or, very roughly:

$$15 \text{ if } P > 100$$

This gives the size of the class-interval as the

$\underline{I} = \sigma^{-1} \underline{I} / \Sigma_i |$  the range divided by the number of intervals (and adjusted to some approximate convenient integral amount, of course).

22. Successive letters of the alphabet are used to denote successive orders of class-intervals and subclass-intervals in a hierarchy,  $i : j : k$ . It is convenient to identify any order of class-intervals by the letter of its order and to use primes to distinguish a particular class within that order. Thus, in  $i : j : kI, j'I, j''I, j'''I, j''''I$ , etc., identify four particular classes of the  $j$  order of subclasses. See S. 5, Ch. II, and S. 1, Ch. III. A skipped or empty class may be termed a "nul" class and is denoted by a script of zero,  ${}_0I$ .

The class-interval script denotes quantitatively equal intervals, or the I-units in which that indicant is measured. Qualitative classes and subclasses are expressed by a hierarchy of class scripts. Accordingly, the use of the class-interval script on the attribute is avoided.

23. These two values of the variable are denoted by exponents of zero for presence ( $I^0 = 1$ ) and minus infinity for absence ( $I^{-\infty} = 0$ ). Since it is a two-point variable it can be more compactly symbolized by a point script of 2, thus  ${}^2I$ .

24. The unit value  $I^0$  times its frequency ( $p$ ) plus the zero value ( $I^{-\infty}$ ) times its frequency is the numerator which, divided by the total frequency, is a proportion. This proportion, multiplied by 100 to shift to "per centum" units, is the percentage.

$$\frac{\sum_1^p I^0 + \sum_1^{P-p} I^{-\infty}}{P} = \frac{\sum_1^p 1 + \sum_1^{P-p} 0}{P} = p/P,$$

$$100p/P = \%P \quad (\text{Eq. 17, Ch. III})$$

25. A rank series is symbolized by:

$$1\text{st, } 2\text{nd, } 3\text{rd, } \dots, N\text{th} = 'I, ''I, ''''I, \dots, NI \quad (\text{Eq. 18a, Ch. III})$$

$$\text{or } 'I + 'd = ''I, \quad ''I + 'd = ''''I, \text{ etc.} \quad (\text{Eq. 18b, Ch. III})$$

where  $d$  is an unknown variable, a first order difference.

Ordinal numbers may be averaged and combined in various ways if their inherent approximateness to cardinal numbers is kept in mind. An average of ordinals is more accurate than any one ordinal alone, since in averaging, the increment between successive ranks becomes the average  $d$  in the equation above, and this best represents all the variable values of  $d$ . As  $d$  approaches a constant increment, the ordinal approaches a cardinal. Alternatively stated, the problem of refining ordinal to cardinal measurements is the problem of reducing the sum of the differences between the  $d$ 's, or the second-order differences ( $d''$ ) of the  $I$ 's, to zero.

$$\sum_1^{N-2} d'' = 0 \quad (\text{Eq. 18c, Ch. III})$$

26. For an excellent elementary exposition of matrix algebra (designed for use in Psychology) the sociologist cannot do better than master the introductory chapter of *The Vectors of Mind* (Ref. 77).

27. Matrices may be of different "degrees." Ordinarily, with rows and columns as shown by 2 multiple descripts, it is a second-degree matrix. A single array (row or column) symbolized by one multiple descript is a first-degree matrix (also called an "array vector"). The single cell, symbolized by the singular descript, may be called a zero-degree matrix. This last is the lower limit of matrices

and is identical with a single index, or indicator. The matrix, from this point of view, is the more inclusive term.

If each cell of a second-degree matrix is subdivided into a sagittal array, it becomes a third-degree matrix. Successive subdivisions of cells can extend it to an  $m$ th degree matrix. The degree of the matrix in an S-formula is the number of multiple aggregative descripts. Thus S. 10, Ch. III is a second-degree matrix, as it has two multiple descripts. S. 43, Ch. XI is a fifth-degree matrix as it has five aggregative multiple descripts.

The *order* of a matrix is the number of cells in each array. Thus, S. 10, Ch. III, is a  $5 \times 47$  order matrix, while S. 43, Ch. XI is a matrix of order  $3 \times 2 \times 2 \times 3 \times 2$  (i.e.,  $\uparrow = 3$  dates,  $\downarrow = 2$  periods,  $\text{p} = 3$  regional plurels,  $\text{q} = 3$  occupational plurels, and  $\text{r} = 2$  racial plurels). A square matrix is of order  $n \times n$ , a rectangular one is of order  $n \times n'$  ( $n \neq n'$ ).

The colon symbol denotes an aggregation. It states that, for each cell of the matrix preceding the colon, there is a further matrix as specified by the multiple descripts on the indices following the colon. Thus, in the Russian 5-Year Plan, S. 31, Ch. II, the three multiple descripts  $\downarrow$ ,  $\downarrow_i$ ,  $\downarrow_j$  denote a third-degree matrix of order  $4 \times 3 \times 2$ . The first colon states that each of the four period cells,  $\downarrow$ , is expanded into arrays of 3 attribute cells,  $\downarrow_i$ , and the second colon further expands each of these 12 attribute cells into an array of 2 further attribute cells,  $\downarrow_j$ , (the "planned" and "fulfilled" categories) resulting in  $4 \times 3 \times 2 = 24$  cells. This matrix is one of one-way dependence, i.e., one index, the annual period, dominates or controls the data. Corresponding to it are three kinds of things observed as freight, passengers, and capital. Dependent on these are the subcategories of "planned" and "fulfilled." The left-hand side of the single colon always dominates the dependent right-hand side. The left side is the observationally independent variable. Whenever there is mutuality and not dominance-dependence the colon is doubled. Thus in S. 7, Ch. II, the correlation scattergram of temperature and national athletic records, the formula is  $\uparrow :: \uparrow : \text{P}_p$ , denoting a mutual cross-classification of the  $\uparrow$  ranks of the temperature indicant with the  $\uparrow$  class-intervals of the athletic indicant into  $i \times j$  cells, and the dependence of the national frequencies,  $\text{p}$ , on each of  $i \times j$  scattergram cells. The colon, in brief, asserts a matrix derived by subclassification; the double colon asserts a matrix derived by cross-classification.

28. Unreliability errors may be lumped together in a total unreliability variance, or isolated as above, or still further analyzed. Thus, observer errors in one study (Ref. 12, pp. 62-73) were isolated and measured as due to differing, (a) questioners, (b) recorders, and (c) scorers of interviews filling out a scored schedule card. Observee unreliability was analyzed and measured separately as due to differing, (a) informants in one family, (b) sex of informant, and (c) date (daily and seasonal intervals) of informing. For formulae for synthesizing these errors see Chapter VI on Correlation. For a formula separating reader error (observer) from content error (observand) of school examinations, see Ref. 27. The term "observation errors" includes the observer, observand, and observee errors.

29. Improvement of measurement means, of course, not only its quantitative

refinement (chiefly reliability) but also improvement in its significance (chiefly validity). Too many studies with reliable technics are made on socially non-significant phenomena, perhaps because a Ph.D. thesis must be ground out and the data are at hand. Unless the criteria for validation are socially significant ones, the validity of an indicator, however high the coefficients, remains a trivial finding. Those who depreciate quantitative methods are justified to the extent that such methods are misapplied or applied to sociologically unimportant problems. Such trivial application however, is equally possible with qualitative technics of any school whatever.

30. This view may help resolve some of the current controversy as to what constitutes "measurement." One study may expressly define the term as beginning with ordinals but excluding all-or-none phenomena. (Compare Yule's view, Ref. 83, Ch. I. His attribute is our attribute; his positive and negative attributes are our all-or-none indicant; and his "variable" is anything from our ordinal up.) Another student may prefer to define measurement as beginning with cardinal units and call our all-or-none and ordinal categories quasi-measurement. In this volume we draw the line between the qualitative and the quantitative, between non-measurement and measurement, between the unitary constant and its variation into degrees, at the point where an attribute becomes an all-or-none variable (a primitive indicant).

An interesting corollary of Yule's terms of "positive and negative attributes" is, that if the absence of the attribute is assigned an arithmetic value of minus one (instead of the usual value of zero), there results a two-category variable, ranging from +1 to -1. If the categories are of equal frequency, the mean is zero and standard deviation and variance are unity. This may have useful implications in interpreting correlation in terms of percentages of such "atomic" elements. (See Ch. VI.)

31. Compare Chapin's levels of symbolic substitution in S. 1, Ch. XII.

32. For some examples of this process in measuring culture complexes, see such scales and their descriptive literature as in Refs. 2, 12, 40, 76.

33. The case script is used chiefly for such points as:

$\bar{I}$  = the mean of the indicant. (See S. 15, Ch. III) (Eq. 24a, Ch. III)

$\bar{I}$  = a two-point indicant, usually an all-or-none indicant (Eq. 16b, Ch. III)

$\bar{I}$  = an indicant of  $i$  points, usually an ordinal indicant. (See S. 7, 24, Ch. II) (Eq. 16c, Ch. III)

$\bar{I}$  = the standard deviation of the indicant, the point which is one sigma from the mean. (Distinguish  $\bar{I}$  = the indicant expressed in sigma units, i.e., divided by sigma.) (See S. 30, vs. 34, Ch. II) (Eq. 24b, Ch. III)

$\bar{I}$  = the lower and corresponding upper limits of the range of the values of the indicant. (See S. 30, Ch. II) (Eq. 24c, Ch. III)

$\bar{I}$  = deviations from the mean;  $\bar{I}$  = a particular deviation (Eq. 24d, Ch. III)

$\bar{I}$  = the zero point or origin. (See S. 9, Ch. III) (Eq. 24e, Ch. III)

It should be noted that summing an aggregative case script (or any aggregative

gative descript) converts attributes into an indicant and a nullary index into a primary index. Suppose the attribute is an obvious unit, such as "a sheep,"  ${}^1I^0$ , then more than one sheep is denotable as  ${}^1I^0$ . This denotes an aggregation, a matrix of the first degree. It connotes, "For each separate sheep . . . some corresponding phenomenon." But as soon as the sheep are summed up in a single number of sheep, then "a sheep" is the unit of an indicant, and the operation of summing the aggregation has converted the attributes into a particular value of an indicant  ${}^1I^{+1}$ . Thus:

$$\Sigma {}^1I^0 = {}^1I^{+1} \quad (\text{Eq. 25a, Ch. III})$$

In another manner the multiple case script may be the equivalent of the singular class-interval script. Thus, a number of individual sheep (cases) may be collectively the class-interval, such as "the third dozen" of sheep:

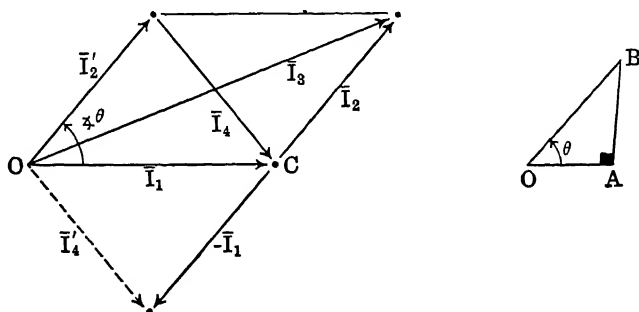
$$\Sigma {}^1I^0 = ., I^{+1} \quad (\text{Eq. 25b, Ch. III})$$

34. A line vector is a line with magnitude and direction, i.e., length and an angle with some line of reference. If overlining denotes a line vector, then an index in vector terms is:

$$\bar{I} = I, \bar{I} \quad (\text{Eq. 26, Ch. III})$$

where the  $I$  not overlined is the magnitude or length, and the  $\bar{I}$  is a "unit vector," a vector whose length is one unit. The  $I$  is called the "scalar" aspect of the vector, or simply a scalar.

A negative vector is opposite in direction to a positive vector. Thus,  $-\bar{I}$ , is negative vector of  $+\bar{I}$ , both of which start at the point  $C$  in the accompanying diagram.



To add vectors they are placed end to end and their sum, called the resultant vector, is the line from the origin to the tip of the last vector. Thus:

$$\bar{I}_1 + \bar{I}_2 = \bar{I}_3 \quad (\text{Eq. 27a, Ch. III})$$

For scalars expressed in sigma units:

$$\sigma_3^2 = \sigma_1^2 + \sigma_2^2 + 2r_{12}\sigma_1\sigma_2 \quad (\text{Eq. 28a, Ch. III})$$

This relation (Eq. 27a, Ch. III) enables the calculation of the length of  $\bar{I}_3$  from a knowledge of the standard deviations of indices 1 and 2, represented by vectors  $\bar{I}_1$  and  $\bar{I}_2$ .

To subtract vectors, add the negative of the vector subtracted, thus:

$$\bar{I}_1 - \bar{I}_2 = \bar{I}_4 \quad (\text{Eq. 27b, Ch. III})$$

Note that, in the accompanying diagram,  $\bar{I}_4'$  is equal and parallel to  $\bar{I}_4$ . The location of the vector is immaterial; it is defined as a line of a specified *length* extended in a specified *direction*.

For scalars in  $\sigma$  units:

$$\sigma_4^2 = \sigma_1^2 + \sigma_2^2 - 2r_{12}\sigma_1\sigma_2 \quad (\text{Eq. 28b, Ch. III})$$

Note that the sum and the difference of two vectors are the diagonals of the parallelogram they make.

There are two ways of multiplying vectors together, only one of which, the "scalar product," is useful in S-theory. (Division by a vector, except in the collinear case, is ambiguous and so is not permitted in vectorial algebra.) The scalar product (also called the dot product) of two vectors is:

$$\bar{I}_1 \cdot \bar{I}_2 = I_1 I_2 \cos \theta \quad (\text{Eq. 29a, Ch. III})$$

which is the product of their scalar aspects times the cosine of the angle theta, which is the angle between the two vectors.

$\cos \theta = OA/OB$  the ratio in the right-angled triangle, in the diagram above, of the base OA, to OB, the hypotenuse (Eq. 30, Ch. III)

Since the correlation coefficient,  $r$ , is geometrically the cosine of the angle between the two vectors which represent the variables correlated, the scalar product can be written:

$$\bar{I}_1 \cdot \bar{I}_2 (= I_1 \bar{I}_1 \cdot I_2 \bar{I}_2) = I_1 I_2 r_{12} \quad (\text{Eq. 29b, Ch. III})$$

Since the correlation is an observable quantity and since the cosine can be converted into degrees of angle from a table, the angle between two components can be graphed geometrically. Thus, S. 34, Ch. II, studies the college admission indicants which will best predict the criterion, college achievement, indicated by graduation standing. The correlations of the various predictive indicants with their criterial indicants are plotted as cosines of angles (in the circle of unit radius where the scalar magnitude of each vector is expressed as one standard deviation). The indicant whose vector makes the smallest angle with the horizontally placed criterion vector has the correlation most nearly perfect, and is, therefore, the best predictor to use as a requirement for admission to this college. The indicant whose vector is most nearly perpendicular to the criterial vector has a correlation most nearly zero and is, therefore, the least efficient in predicting academic success in this college.

The square of a vector is the square of its scalar,

$$\bar{I}^2 = \bar{I} \cdot \bar{I} = II \cos \theta = IIr = III = I^2 \quad (\text{Eq. 31, Ch. III})$$

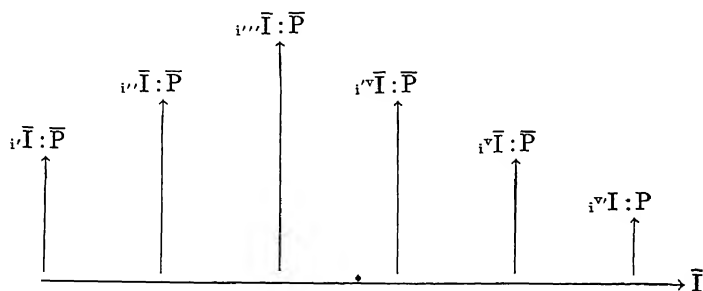
for the vector makes an angle of zero with itself and the cosine of an angle of zero degrees is unity (i.e., the index correlates perfectly with itself so that  $r = 1.00$ ).

35. The lower limit of the matrix (as previously stated) is reached when it shrinks down to a single cell and becomes identical with an index. An index is

a single resultant vector no matter how many components may enter as factors or addends into that index.

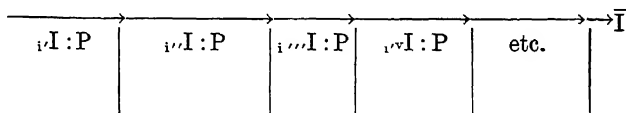
36. For another example of the relations of indices, vectors, and matrix notation consider S. 10, Ch. III, with the formula,  $I_i:j$ . This is a second-degree matrix, as shown by two multiple descripts,  $|_i$  and  $|_j$ , of the  $47 \times 5$  order as shown by the legend, where  $|_i = 47$  environmental items, and  $|_j = 5$  quintile probability indicants. There are  $47 + 5 = 52$  vectors representing this situation. In the 5-dimensional space determined by the quintile indicants as reference vectors, or co-ordinates, 47 resultant vectors are defined. Each of these 47 row vectors is defined by the five cell entries giving the scalar distance along each of the five reference vectors. These five distances determine a point in 5-space which, together with the origin, determines a row vector representing that environmental item. Alternatively, the reference and resultant vectors may be exchanged, letting the 47 vectors be considered as co-ordinates, or reference vectors, defining a space of 47 dimensions. It may be an oblique space if the environmental items are correlated between 0 or 1.00, showing cosines of oblique angles between the 47 vectors. In this 47-space the five quintile column vectors can be considered as resultant vectors. Each column vector is determined by the origin point and a second point specified by the 47 cell entries which state 47 distances along the 47 reference (i.e., row) vectors respectively.

37. One method of graphing the aggregative operation of the colon is in the form of the teeth of a rake:



where the "teeth" may be of uneven length. Thus, in the formula for a frequency distribution,  $iI:P$ , the indicant is the horizontal vector from each of the line-sects (class-intervals,  $i'$ ,  $i''$ ,  $i'''$ , etc.) from which springs a dependent  $\bar{P}$  vector.

A second geometric interpretational hypothesis is to consider the dependent index ( $P$  above) as a scalar whose corresponding values are weights lengthening the class-interval of the independent vector,  $I$ , thus:



This interpretation makes the colon the symbol for a distribution product of a vector and a scalar quantity—a form of multiplication intermediate between

the ordinary arithmetic multiplication of two scalar quantities and vectorial multiplication of two vectors (in the dot product described above).

It should be noted that the one-way dependence, the amount of which is summarized in one of the two correlation ratios, is represented by the colon. Two-way, or mutual dependence, is where the correlation is rectilinear, and the two correlation ratios become the one correlation coefficient and are represented in S-theory by the double colon (if the operation is carried only as far as cross-classifying in a scattergram), or by the heavy dot (if the operation is carried through to calculate the correlation coefficient, which is the scalar product of unit vectors). From this viewpoint the colon denotes an imperfect, or half-way stage of vectorial multiplication. In terms of the scalar product as a cosine of the angle between two vectors,  $I$  and  $J$ , the colon suggests a situation where the angle from  $\bar{I}$  to  $\bar{J}$  might be different, and even have no constant relation to the angle from  $\bar{J}$  to  $\bar{I}$ . This irreversibility occurs again in the interrelation matrices of Chapter VII where the "social distance" of party  $P$ , to party  $P_{,,}$  may be unequal to the social distance of  $P_{,,}$  towards  $P$ , (i.e., the "social distance margin" =  $P, : P_{,,} : I - P_{,,} : P, : I \neq 0$ , or in Brief-S formula,  $I, : , , , : ,$  as explained in Chapter IX).

In sum, the colon can be conceived as an operational variable, varying over three values. These are three kinds of multiplicative operations which form a series as follows:

1. Colon absent =  $II = II$  = scalar times scalar = ordinary multiplication  
(Eq. 33a, Ch. III)
2. Colon present =  $I : I^? = \bar{I}I$  = vector times scalar = semi-vectorial multiplication  
(Eq. 33b, Ch. III)
3. Double colon =  $I :: I = \bar{I} \cdot \bar{I}$  = vector times vector = scalar product of vectors  
(Eq. 33c, Ch. III)

Our colon notation may become dispensable in developing vector notation to cover qualitative data as in attributes ( $I^0$ ), one-way dependency as in the correlation ratios and incompletely computed situations. At present, the colon seems a more flexible symbol; but since all of S-theory, especially its operational symbols, is hypotheses, further data and the criticism of other scientists may be expected to lead to their modification just as well as to their rejection or adoption.

The division sign,  $/$ , denotes scalar division. Division of a vector must always be by a scalar quantity, since a vector as divisor gives ambiguous results and is ruled out by mathematicians.

Division by a vector is possible in the case of collinear vectors which are vectors that are parallel. Tensor theory develops a form of vectorial division by multiplying an inverse in the  $i, j, k$ , system, but the application of this to S-theory remains to be explored.

Thus, any index with a negative exponent is considered as a scalar quantity. In dividing a vector by a scalar quantity the scalar aspect of the vector is divided by the divisor scalar, resulting in a vector unchanged except as to length. Thus:

$$\bar{I}/I_{,,} = I\bar{I}/I_{,,} = (I/I_{,,})\bar{I}, \quad (\text{Eq. 34, Ch. III})$$

The addition and subtraction of vectors ( $\overline{+}$   $\overline{-}$ ) is explained above. The multiplication sign ( $\times$ ) denotes another kind of product of vectors, called the "vector product" by mathematicians to distinguish it from the "scalar product of vectors" denoted by the heavy dot. This vector product (see S. 14, Ch. VI) is still unexplored as to its usefulness in S-theory.

38. The adjective "apparent" is inserted, as the  $s$  vectors may be found to lie in a space of less than  $s$  dimensions, as when three vectors lie in one plane. The number of necessary dimensions is set by the "rank" of the matrix of the inter-correlation coefficients of the indices. This bit of the theory of factors (or of components as they are beginning to be called in psychological usage) is more fully discussed in Chapter VI on Correlation. For an excellent exposition introducing the beginner, see Thurstone, L. L., *Vectors of Mind*, Univ. of Chicago Press, 1935, p. 266. This field growing out of C. Spearman's two-factor theory, and G. H. Thomson's sampling theory, and greatly enlarged by Thurstone, Kelley, Hotelling, and others is rapidly growing as evidenced by the new journal *Psychometrika*.

39. In the case of the space sector there are two types of dimensions (as developed more fully in Chs. VIII and X), the ordinary three physical dimensions connoted by "length," "breadth," and "height," and the statistical dimensions specified by the statistical moments (Eqs. 7-11, Ch. V). The number of physical dimensions is given by the spatial exponent ( $|$ ), and this times the number of statistical dimensions gives the total number of spatial dimensions in a situation.

40. Non-orthogonal, or "oblique," dimensions may be used. An oblique set of dimensions can be transformed into an orthogonal set, and vice versa, by appropriate transformations. These transformational technics involve the matrices of the intercorrelation coefficients as these specify the angle between every pair of components.

41. It should be noted that the term "dimensions" is ambiguous in mathematics and science. One has to specify the "dimensions of physics," "of algebra," "of Euclidean geometry," or "of some other geometry." Thus, for one lack of consistency, a zero-dimensional space in geometry is a point which is also equivalent to a line of zero length. But in algebra, the zero exponent reduces its base letter to a value of unity. Thus,  $x^0 = 1$ , and in co-ordinate geometry, which links algebra and geometry, this means an  $x$  co-ordinate of unit length, not a dimensionless point.

In algebra, "dimensions" connote the degree of the equation as given by the exponents. In geometry, "dimensions" connote the number of variables in an expression, i.e., the number of Cartesian co-ordinates. Thus:

$ax + b = 0$  is a first degree (linear) equation involving one co-ordinate  
(1 variable) (Eq. 36a, Ch. III)

$ax + by = 0$  is a first degree (linear) equation involving two co-ordinates  
(2 variables) (Eq. 36b, Ch. III)

$ax^2 + by = 0$  is a second degree (quadratic) equation involving two co-ordinates (2 variables) (Eq. 36c, Ch. III)

$ax + by + cz = 0$  is a first degree (linear) equation involving three co-ordinates  
(3 variables) (Eq. 36d, Ch. III)

In Physics, the "dimensions" mean the exponents on M, T, and L, the symbols for mass, time, and length. Thus, the dimensional formula for a velocity is  $LT^{-1}$ , and for a physical force,  $MLT^{-2}$ .

In S-theory the number of dimensions is determined by the class script. The class script states the number of indices which are variables in the geometric sense, and it connotes the number of Cartesian co-ordinates required to geometrically represent that S-situation. The exponent deals with the directions of the dimensions—an exponent of zero in general fixing a direction by asserting a point in societal multidimensional S-space, an exponent of +1 asserting a distance in that direction, and an exponent of +2 asserting that the angle between two dimensions has been determined. The exponent thus states the operational degree to which the directions of the dimensions have been determined. (The spatial sector is complicated by the possibility of a double operation of exponents, first in squaring or cubing a length, and second in calculating the statistical moments (as sketched in the previous sentence) in frequency distributions of such lines, areas, or volumes.)

# *PART III*

## THE POPULATION SECTOR, $\mathbf{P}_p \mathbf{P}_p$

*studying situations defined by  $S = T^0; I^i; L^0; P^{1,2}$*



## Chapter IV

### PLURELS, P

#### I. THE POPULATIONAL EXPONENT, $|^p$

##### A. Definitions

A single human being will be referred to as a "person." Any collection or grouping of persons will be defined as a "plurel." A group is an interacting plurel. A population means all the persons and plurels within it. A "party" is the term which will hereafter denote *either* a person or a plurel. A "population" may vary upwards in number from one party at the lower limit. Occasionally the more colloquial term "people" will be used as practically synonymous with population. A population index is the size of a population stated either in relative terms, such as percentages, or in absolute terms of the number of parties.

Population indices denote populations that are qualitatively different, i.e., are products of different attributes and population. A population index is thus not entirely synonymous with the term plurel, for in a frequency distribution the frequencies of persons corresponding to each class-interval of the indicant are a plurel, but only a part of the population. Such a plurel is a quantitatively defined plurel having the same attribute as the plurels in the other class-intervals. Such a plurel geometrically is a line-sect, a scalar subdivision of the length of the vector which represents the index.

All the population indices, modified by their scripts, constitute the population "sector" of any defined situation.

##### B. The Populational Exponent

In societal situations population indices may occur with any one of four exponents from minus one to plus two:

$$P^p = P^0, P^{+1}, P^{-1}, P^{+2} \quad (\text{Eq. 1a, Ch. IV})$$

including a few combinations of them.

1. NULLARY POPULATIONS,  $P^0$ 

Whenever a situation has little or no reference to the number of persons, the population index is "nul" and is so denoted by an exponent of zero and no descript.

$P_0^0$  = a "nul" population, not affecting the situation (an index with nullary exponent and zero descript)

(Eq. 1b, Ch. IV)

All the S-situations in Chapter III on Indicators contained nul populations, which, since they had zero descripts, were not written in the descriptive formulae and were denoted in the quantic number by a zero. Of course, these and all other societal situations to be *societal* ones must have some reference, however remote, to people. Either they are created by people as part of human culture, or at least, to be recorded and presented as an S-situation, must have been observed by people. But if the *number* of persons is immaterial in the situation, if a canvass or enumeration of persons or plurels has not taken place, either actually or theoretically, in the recorded situation as presented, then the population is defined as "nul."

2. PRIMARY POPULATIONS,  $P^1$ 

The operation of listing or counting persons or plurels defines the existence of the primary population index denoted by an exponent of plus one:

$P^{+1}$  = a population, a primary populational index

(Eq. 1c, Ch. IV)

All the S-situations in this chapter are examples of primary populational indices.

Whenever the population is expressed, not in absolute units of the number of parties, but in relative units of percents or per-milles, one population has been divided by another population. This merely shifts the units.<sup>1 \*</sup>

3. SECONDARY POPULATIONS,  $P^2$ 

After a population, characterized in some way, has been observed, a more penetrating level of sociological observation is to

\* For Eqs. 1e-2e, Ch. IV, see notes at end of the chapter.

determine its structure, and the interrelations of the persons and plurels composing it. The interrelation of every person with every person, including self-relations, and the interrelations of every plurel with every plurel, including in-group relations, are seen in S-theory to be the core of distinctively sociological phenomena as distinguished from the phenomena of all the other social sciences which S-theory also covers. (It is to study these interpersonal relations more intensively that the new journal, *Sociometry*, has been founded.) These interrelations, including relations between persons and plurels, are most systematically set forth in the *population interrelation matrix*. This matrix has a column for each party, i.e., person or plurel, and a row for each party, so that there is a cell (at the intersection of a row and column) for every possible pair of parties. If P is the number of parties, the complete interrelation matrix has  $P^2$  cells. These and other considerations (which are more fully presented in Chapter VII on Interrelations) make the exponent of 2 on the population index the quantic formula, symbolizing the interrelations of the parties in a population.

$P^2$  = the quantic formula for an interrelated population, a group, a matrix of the interrelations of all pairs of persons and plurels in a population (Eq. 1d, Ch. IV)

(For examples see all the S-situations of Chapter VII.)

### C. Plurels in the Quantic Classification

The phenomena of interrelated populations ( $P^2$ ) will be taken up in Chapter VII. The present Chapter IV deals only with primary populations,  $P^{+1}$  and  $P^{-1}$ . In the three-dimensional model of the quantic solid (shown in S. 33, Ch. II) the previous Part II and Chapter III on Indicators studied the bottom stratum,  $P^0$ , where population was nul, and in it the array from the leftmost to the foremost corners where time was nul,  $T^0$ , i.e., static situations. The present Part III on the Population Sector takes up first in Chapters IV, V, and VI the array in the middle stratum, defined by  $P^1$ , ( $|^s = 0; i; 0; 1$ ) under the Indicator array studied above, and next in Chapter VII the array in the top stratum defined by  $P^2$  ( $0; i; 0; 2$ ). The present chapter takes up one cell of the  $T^0 I^1 L^0 P^1$  array ( $0; i; 0; 1$ ), namely the cell of static attributes

combined with population and defined by the quantic formula,  $T^0;I^0;L^0;P^1$ , or by the quantic number  $0;0;0;1$ . This is the cell of persons and plurels, which are qualitatively characterized without reference to lapse of time or to geographic area.

Population characterized by quantitative indicators, i.e., measured by some indicant, will be studied in the next chapter on Distributions. Following this, Chapter VI on Correlation studies populations characterized by correlated indicators.

Populations distributed with reference to geographic area will be studied under the heading of Densities in Chapter VIII, when the spatial indices will be combined with indicators and populations. Populations distributed in time, as in age pyramids, vital statistics rates, and other populational changes, will be taken up in Chapters IX to XI of Part V, where temporal indices will be combined with the indicatory, populational, and spatial indices.

## II. THE PLUREL SCRIPT, $|_p$

### A. Definition of the Class Script, the Post-Subscript, $|_p$

The number and the kind of plurels which may compose a population are denoted by the class script, the post-subscript,  $|_p$ . The class script in the population sector is conveniently referred to as the "plurel script." As usual, the primed script,  $|_p'$  (or the prime alone), called the "singular" plurel script, denotes a particular specified plurel; while the unprimed lower case letter,  $|_p$ , called the "multiple" plurel script, denotes the aggregation of plurels of a given kind. The legend states the kind of plurel in the given S-situation that is symbolized by the plurel script. A population is usually composed of a hierarchy of plurels, that is, a general classification with classes and subclasses and sub-subclasses in descending "degrees":

$P_{p':q:r}$  = a hierarchy of plurels showing 3 degrees of plurel scripts  
(Eq. 3a, Ch. IV)

This formula states that there is a population of persons,  $P$  in number, of the kind denoted by  $|_p'$ , which has corresponding sub-plurels of the kind and number denoted by  $|_q$ , each of which in turn is subdivided into subplurels of the kind and number denoted by  $|_r$ , each of which in turn might be composed of further subplurels of the kind and number denoted by  $|_s$ . The colon

symbol, denoting correspondence in general, is more simply verbalized here as subclassification.  $|_{q:r}$  may be read, for example, as "the q plurels are each subclassified into a variable number of r plurels." In matrix terms, a cell is extended into an array by being broken down into a series of subcells. The organization chart of a business firm, a government department, or region, into state, county, and township strata are examples of a hierarchy of plurels. This hierarchy can be visualized in the diagrammatic pattern:

$$P_s = \left\{ \begin{array}{l} \text{1st degree, } |_{p'} \longrightarrow |_{p'} \\ \text{2nd degree, } |_q \longrightarrow \begin{array}{|c|c|c|c|} \hline |_{q'} & |_{q''} & |_{q'''} & |_{q^{iv}} \\ \hline \end{array} \\ \text{3rd degree, } |_r \longrightarrow \begin{array}{|c|c|c|c|c|c|c|c|} \hline |_{q'r'} & |_{q'r''} & |_{q'r'''} & |_{q'r^{iv}} & |_{q''r'} & |_{q''r''} & |_{q''r'''} & |_{q''r^{iv}} & |_{q'''r'} & |_{q'''r''} & |_{q'''r'''} & |_{q'''r^{iv}} & |_{q^{iv}r'} & |_{q^{iv}r''} & |_{q^{iv}r'''} & |_{q^{iv}r^{iv}} \\ \hline \end{array} \end{array} \right\} = P_{p':q:r} \quad (\text{Eq. 3b, Ch. IV})$$

Note should be made of the fact that this hierarchy of classes can be general to all four sectors and becomes a hierarchy of plurels only in the population sector. The number of subdivisions is a variable. The variable may include the value of zero which denotes a "nul" class. Thus in the diagram above, the four plurels in the second degree, identified by  $|_{q'}$ ,  $|_{q''}$ ,  $|_{q'''}$ ,  $|_{q^{iv}}$ , have 4, 2, 3, and 0 subplurels respectively in the r degree or stratum.  $|_r$  then, has four values here, namely,  $|_r = 4, 2, 3, 0$ . A further important property is that the classes are non-overlapping. (The case of overlapping plurels with common members is operationally more complex and is dealt with in Chapter VII.)

It may readily be seen that this hierarchy formula (Eq. 3, Ch. IV) has a thousand applications, family trees, tournament charts, army ranks, jurisdictions of judicial courts, electoral districts, precincts and wards, departments of a factory or store, or any organizational scheme, or any classification of people into classes and subclasses. The reader should scan S-situations, in this and other chapters, which involve a hierarchy of class scripts for examples of its significance.

### *B. The Product of an Attribute and a Pure Population, the Described P*

A closer scrutiny of the meaning of the plurel script can now be made. In reality the plurel script represents an implicit

product of an attribute with a class script multiplied by a "pure" population. In this statement the terms "implicit," "product," and "pure" should be clearly understood. A "pure" population is a *number of persons* with no indication of any sort as to the kind, identity, or characteristics of those persons. The qualitative characteristics which state the kind and identity of the population are expressed by an attribute with its descripts. Thus an enumeration of the population of France is in reality a product of the characteristic "French" =  $I_i^0$ , times a number of persons,  $P$ , so characterized. For simplicity this product is cross-scripted,<sup>2</sup> i.e., condensed in notation, so that the plurel script denotes what the attribute class script would denote:

$$I_i^0 \cdot P_0^{+1} = P_p^{+1} = \text{the product of a population and the attribute characterizing it} \quad (\text{Eq. 4a, Ch. IV})$$

The "pure  $P$ " in the left-hand member of Eq. 4, Ch. IV is a mere coefficient. Its zero plurel script states that it is a plurel of no *kind* whatever. Vectorially considered, it is a scalar length of the vector whose direction is defined by the qualitative characteristic, the attribute. The "descripted  $P$ ," defined by the right-hand member of Eq. 4, Ch. IV is an implicit, or understood, product of an attribute and a "pure" population. Geometrically, it is a vector whose scalar length is stated as the number of its person-units,  $P$ , and whose direction is stated by the kind of population denoted by the plurel script,  $|_p$ .

This concept of a "pure" person and a "pure" population derives from our definition of a person as a human being having characteristics.<sup>3</sup> Every pure person is considered as equal and interchangeable with every other pure person. It is a standardized cardinal unit of population which is objectively determinable as in "counting noses." All persons are identical in the one respect of their being carriers of characteristics. These characteristics, qualitative and quantitative, static and dynamic, in their myriad patterns, distinguish persons from each other yielding the observed individual differences of behavior and personality. The pure person is an abstract concept created for the purpose of standardizing a unit of population. No one ever saw or touched a pure person apart from any and all his characteristics. A pure

person is defined as a single human being, abstracted apart from all his characteristics.

The ways in which observed persons differ from each other define their variable characteristics. This definition is clear-cut and unambiguous, as persons, even in the case of Siamese twins, can be identified by counting heads and *all* other differences and properties are characteristics. The problem of predicting and controlling human behavior is thus chiefly a problem, as stated before, of developing *indicators* of the characteristics of people (in suitable combinations with the time and space indices, of course).

A "pure population" is a number of "pure persons," unidentified and uncharacterized in any way except as to number. It is thus readily seen that *every* observed population is a combination of some characteristic observed as an indicator and a pure population. This combination is a product. It can be represented mathematically as an attribute times a pure population, a quality times a quantity, yielding as product a quantity of a quality, an amount of some kind of people. This statement is symbolically stated in Eq. 4, Ch. IV. This attribute-population product is a fundamental concept in the S-theory, as is the product of the attribute with any index. As noted above in Eq. 9, Ch. III, every indicant is implicitly such a product of an attribute and a numerical coefficient, a "pure" indicant. As will be seen in later chapters all identified geographic regions and all dated historical events are similarly products of attributes with pure, i.e., mathematical, space and time. The *attributes* represent any item of qualitative knowledge. Its product with quantities is a major hypothesis in S-theory and may be referred to as the "attribute product" hypothesis. Its usefulness in combining qualitative and quantitative phenomena within the field of mathematical manipulation is nowhere better illustrated than in the population sector.

The attribute product can next be extended to the multiple plurel script, thus:

$$I_i^0 P_0^{+1} = P_p = \text{a population composed of plurels of the kind and number denoted by } |_p, \text{ each plurel having the number of persons denoted by the (variable) } P$$

(Eq. 4b, Ch. IV)

Note that when  $P$  is alone, a plurel script of 1 is always understood. It denotes the number of persons in the whole population in the situation that is being analyzed. When the  $P$  is modified by the plurel script it becomes a variable in the form of a matrix with as many cells, or numerical values of  $P$ , as there are plurels. Thus in S. 2, Ch. IV,  $P_p$  has three values, namely,  $P_{p'} = 77,118$  "deaths,"  $P_{p''} = 83,390$  "severely wounded," and  $P_{p'''} = 142,104$  "other casualties," while in the subplurels, classified by causes of death,  $P_q$  has four values, namely,  $P_{q'} = 34,248$  "killed in action,"  $P_{q''} = 23,430$  "died of disease,"  $P_{q'''} = 13,700$  "died of wounds," and  $P_{q^{iv}} = 5740$  "died from miscellaneous causes." In each instance  $P$  is the number of persons, and the singular plurel scripts (differentiated within each degree of the classification by the primes) denote the kind of persons that are specified by the phrase in quotation marks. This illustration extends the attribute product to a hierarchy of two degrees.<sup>4\*</sup>

This can be generalized as follows:

$$I^0_{i \cdot j \cdot k \cdot \dots \cdot n} P_0^{+1} = P_{p:q:r:\dots:z} = \text{a hierarchy of plurels defined as the product of a pure population index and a hierarchy of attributes characterizing that population} \quad (\text{Eq. 4c, Ch. IV})$$

Note that in the descriptive formula (as distinguished from the quantic formula which omits the descripts) a  $P$  with zero descripts is a "pure"  $P$ , i.e., an undescripted or undescribed population. Since the pure  $P$  never has any descripts, it is convenient to use the notational condensation of letting the population descripts,  $\mathbb{P}_{\mathbb{P}}$ , denote what is denoted by the descripts of the attribute,  $I^0_{i \cdot j \cdot k \cdot \dots \cdot n}$ , that are multiplied into the  $P$ . This condensed notation saves wearisome redundancy of writing an  $I^0$  diphthonged with every index. It is as justifiable a convention as the omission of the (understood) exponent of 1 and coefficient of 1 attached to every algebraic letter, as in the expression  $x + y - z = 0$ , which really means  $1x^{+1} + 1y^{+1} - 1z^{+1} = 0$ . In practice, in analyzing situations into a descriptive S-formula, the populational descripts are used with little awareness of this implicit condensation. But the student should clearly realize that

\* For Eqs. 5a-b, Ch. IV, see notes at end of the chapter.

every described population represents an attribute-population product, or else the definitions of the sectors become confused and the concepts of this S-system would cease to rigorously fulfill the canon of classification that categories must not overlap and must have definite boundaries.

The product of an attribute and any index is comparable to a noun modified by an adjective (or, in the case of an attribute times a dynamic index ( $IT^{-1}$ ), it is comparable to a verb modified by an adverb). The quality expressed by the attribute limits and modifies the pure index much as an adjective does.<sup>5</sup> \* For example, in S. 10, Ch. IV the formula with the attribute written explicitly is:  $T^0 : I_{p,q,r,s}^0 P$  in which

$I_p^0$  = "gainfully occupied in the United States" (really 3 attributes, i.e., 3 adjectives kaleidoscoped)

$I_q^0$  = "male" and "female"

$I_r^0$  = "professional," etc.; "skilled," etc., and "unskilled"

$I_s^0$  = "clerical," "agricultural," "commercial," etc.

and  $P$  = the number of persons so characterized

The question arises as to when, if ever, the attribute is written explicitly with the population. Either the explicit or the implicit treatment is defensible usage and makes no difference in the quantic formula classifying the situation. It is partly a matter of emphasis. One attribute is always treated implicitly to identify the population, a second and further attributes are preferably condensed, though occasionally they are written explicitly for separate emphasis (as the statements of beliefs in S. 12, Ch. IV).<sup>6</sup>

### *C. Classification of Plurels*

A plurel has been shown to be a number of people characterized by indicators in the form of an attribute-population product. Can these attributes defining plurels be reduced, for systematic purposes, to a more significant classification than the unsatisfactory general classification of indicators outlined in Chapter III? Eubank has collected together the attempts that sociologists have made to classify plurels (Chapter VIII of Ref. 25). The student should consult this excellent review of the literature, as only a few points from it will be taken up here.

\* For Eq. 6, Ch. IV, see notes at end of the chapter.

Eubank lists the classifications (Ref. 25, pp. 116 ff. and pp. 154 ff.) which have been offered, built upon the following bases (characteristics):

Concepts pertaining to various classifications of plurels—

1. Ethno-anthropological classifications
2. General social classifications
3. Classifications based on culture levels
4. Classifications based on structure
5. Classifications based on function
6. Classifications based on extent of the social contact
7. Classifications based on the nature of the bond that holds the group together
8. Special plurels of significance to Sociology <sup>7</sup>

These classifications Eubank, following Von Wiese, reduces to three types:

categories—plurels having some characteristic in common, i.e., red hair, superior IQ, male sex, a certain age, laborers, criminals, the elite, etc. A classification on the basis of similarity

aggregations—crowds, mobs, people-on-the-street, audiences.

A classification on the basis of spatial proximity

groups—interacting people, a plurel of persons in active or suspended psychic interaction. A classification on the basis of interaction

Of these types, only the qualitative subdivision of the “categories type” would be classified as in the present chapter, by the quantic formula which is defined by  $P$ , i.e.,  $I^0P$  ( $|^s = 0;0;0;1$ ).

The quantitative “categories” are classified by the quantic formula under distributions  $I^{+1};P^{+1}$  ( $|^s = 0;1;0;1$ ), which is the theme of the next chapter. The crowds involve physical space and will therefore be taken up as classified by their quantic formula of  $I;L^2;P$  ( $|^s = 0;1;2;1$ ) in the chapter on Densities. (Part IV, The Space Sector.)

The interactivities, which are “groups” in the newer sociological sense, have the quantic formula of  $I;P^2$  ( $|^s = 0;1;0;2$ ) and are the subject of a special chapter on Interrelations of people. Their definition and classifications of structures and functionings

will be reserved for fuller treatment in Chapter VII (and Chapters X and XI for their dynamic aspects).<sup>8</sup>

For many practical purposes the classification of indicators of group-culture, based on Bernard's "composite or derived institutionalized environments organized for the purposes of social control" (see Chapter III), is a useful though rough classification both of groups (interacting plurels) and of plurels generally. Our elaboration of this list is:

- A. Segmental—plurels specializing on a segment of the life of a person and of a plurel
  - 1. Domestic, including sex plurels
  - 2. Economic, including occupational plurels
  - 3. Political plurels
  - 4. Educational plurels
  - 5. Medical (health) plurels
  - 6. Recreational plurels
  - 7. Ethico-religious plurels
  - 8. Aesthetic plurels
  - 9. Linguistic plurels
  - 10. Racial plurels
- B. Communal—plurels combining in varying degrees the segments above

More or less self-sufficient, or partially isolated, plurels (in which the cleavage between subclasses of one segment, such as economic levels or races, often dominates): gold coast vs. slum communities, nomadic vs. settled communities, resident or native vs. foreign or transient communities, racial-linguistic-religious communities (such as Jews and Gentiles in some countries, or Moslem, and Hindu and English in India), negro vs. white communities in South Africa and in the southern United States, the "intelligentsia" vs. illiterate communities, the "underworld" vs. law-abiding communities, etc.

As previously emphasized, this segmental classification of plurels is open to many criticisms and is far from being a logically satisfactory classification. Its chief value is that it is in practical everyday use in society. All persons when occupationally classified in a census tend to be put under some such set of headings. In totalitarian plurels approaching inclusion of all plurels in a

region, such as government ministries or cabinet departments, there is a similar rough tendency to follow this classification in ministries of economics or commerce, education, health, communications, welfare, and subdivisions of the political into judicial, military, diplomatic, and other departments. In institutions representing a comprehensive classification of knowledge, such as the larger universities, there are departments of sociology and social work, economics and business, political science and law, psychology and education, religion and ethics, languages and literature, fine arts, extra-curricular activities, etc. These roughly correspond to the segmental groupings above, though historical accidents of the institution's development and other factors give rise, of course, to much variation in grouping, subdivision, and emphasis of these fields. (The racial category seems less co-ordinate than the others, but recalling that these categories are of plurels "organized for purposes of social control," and recalling this recent role of race in white vs. black, white vs. yellow, "the white man's burden," Nordic and Aryan and Semitic ideologies, its inclusion is perhaps justified.)

These plurels are to some extent defined by the persons who are specialists in carrying on the function of that plurel on behalf of some clientele or public. Thus housewives and domestic servants specialize in maintaining domestic plurels (families); all gainfully occupied producers and distributors specialize in some kind of goods or services; government and party officials and employees specialize in political functions; teachers and researchers specialize in imparting or gathering knowledge for clienteles of students, scholars, and scientists; the clergy specializes in the functions of religious and ethical clienteles; the doctors, nurses, pharmacists, and the rest of the medical profession, and employees of institutions specializing in purveying health constitute the health specialists of society; similarly, directors and employees of amusement, sporting, fraternal, recreational, and social organizations are the recreational specialists serving this segment of the life of society; teachers of language, interpreters, writers, workers, in publishing and broadcasting establishments specialize in communication functions; and propagandists of inter-racial attitudes specialize in making the characteristics of race a means of societal control of some plurels by others.

The reader of the above sketch is as likely to be as much impressed by the lack of boundaries, the merging or overlapping of one category and another, as by a clearer concept of separate plurels. This is inevitable in society. The patterns of societal life are so interwoven and integrated that any classification of them and of the persons embodying them into segments can be classified only roughly into "regions" with border zones always present. The scientific solution of this problem lies not in facile verbalizing or naming plurels and their characteristics, such as in the above paragraphs, but in the arduous and skilled work of observing these characteristics more minutely, comprehensively, and reliably with the aid of such instruments as the schedule card, attitude test, records of speech, and other behavior, photographs of material culture, and movies of behavior, etc. With the painstaking work of thousands of investigators over scores of years, inventing, validating, applying, and analyzing the findings of instruments of precision for societal observation, the data about plurels and groups will become an increasingly adequate basis for significant classification such as enables prediction and control.

This working hypothesis that better instruments of observation and resultant better data are the paramount needs for developing a real science of society, may be illustrated in the case of the "communal" plurels listed above. To fix thinking, consider the Moslem, Hindu, and English communities in India as an example. What are the limits defining each of these communities? They shade off into Eurasian hybrids in one direction, into outcast Hindus trying to identify themselves with the Sikh or any other community in another direction, into Buddhist, Christian, and other non-Moslem, non-Hindu Indians in another direction, and so on ad infinitum. To what extent does the classifying of people in India into Moslem, Hindu, English, and other communities serve the scientific purposes of improving our ability to predict and control the behavior of such people? Only as knowledge of the habitual action patterns of each community in defined situations becomes more precise and comprehensive can prediction become more accurate, and agencies of purposeful control become more effective. Some obvious phenomena are known. If Moslems kill a cow, sacred to Hindus, before a crowded

Hindu temple, it is a provocative act which may be predicted as tending towards a riot, and police or an influential leader must be promptly called out to control the mob. But subtler phenomena elude the sociologist until equipped with better instruments of observation. How will each community react to such-and-such legislation, or speech, or administrative act? If it were possible to have every person in an audience, or listener to a radio broadcast, or reader of a newspaper, by some electrical system press a button and record his vote or attitude in suitable categories more significant than mere "aye" and "nay," a more sensitive "barometer," or predictor, of public opinion and of emotions and eventual behavior would be achieved. This is but one possibility among a host of instruments of societal observation which must be developed in order to secure more reliable and valid data towards improving societal prediction and control.

The above digression on methodology in classifying plurels leads to another point on methodology. This is that the contribution of S-theory is seen not as an improved classification of plurels, but as a symbolic notation for handling any classification that may be proposed. The classification into segmental and communal plurels above is unsatisfactory and needs to be improved upon. But more useful than the offering of one more classification of plurels to the sociological repertoire may be the formula for a hierarchy of plurels ( $P_{p:q:r:\dots:z}$ ), which are merely characterized ( $P^{+1}$ ), or found in interacting relationships ( $P^{+2}$ ), with some specification of internal structure of leaders or specialists ( $P^{|}$ ), in terms of standardized units ( $_{\mathbb{P}}$ ). This symbolism is summarized as  $\mathbb{P}\mathbb{P}$ :

$$\left. \begin{array}{l} \text{the list of} \\ \text{specified persons} \end{array} \right\} = \mathbb{P} \mathbb{P} = \begin{array}{l} \text{the exponent} \\ P = \text{the number of persons} \end{array}$$

$$\left. \begin{array}{l} \text{the list of} \\ \text{class-intervals} \end{array} \right\} = \mathbb{P} \mathbb{P} : q : r : \dots : z = \begin{array}{l} \text{the list of plurels of successive} \\ \text{degrees in a hierarchy} \\ = \text{the population sector, a hierarchy of persons and plurels,} \\ \text{characterized in defined units} \end{array} \quad (\text{Eq. 7, Ch. IV})$$

This formula enables the multiplying of persons and plurels by their qualitative and quantitative characteristics and rela-

tionships and furthers the precision of thinking required by the mathematical rules of manipulating equations.

#### *D. Sociation, $|_p$*

The plurel script states the number of plurels into which a given population is subdivided. If these plurels (a) include the whole population, (b) without overlap, and (c) are of interacting parties, i.e., are "groups," then the plurel script measures the amount of association or dissociation in that population. The limits of this measure are unity at the extreme of association in one big group, and the number of persons (P) at the other extreme of complete dissociation from groups into single persons.

$|_{\Sigma p}$  = the number of plurels, a measure of sociation (Eq. 8a, Ch. IV)

$|_{\Sigma p} = 1$  = associative limit of sociation (Eq. 8b, Ch. IV)

$|_{\Sigma p} = P$  = dissociative limit of sociation (Eq. 8c, Ch. IV)

Thus, if several religious denominations unite, if several business firms merge, if political states federate together, if two clubs reorganize into one club, or if ten schools teach the pupils formerly taught by more schools, association has taken place. Oppositely, the splitting of a church into two churches, the proliferating of business, political, social, or other organizations into more numerous organizations, are examples of dissociation, provided they comprise the same population and are not formed by recruiting outside of that population. (The *processes* of associating and dissociating are discussed in Chapter X on Change.)

Since the symbol  $|_p$  means plurels, the above formulae include not only sociation of groups but also non-interacting plurels. The latter type is illustrated when the observer differentiates the population into more subclasses, as when a census starts to specify different kinds of "engineers," "doctors," etc. This type of dividing up into more numerous plurels is done by the analyst, not by the people in the plurels. It is not ordinarily thought of by sociologists as association or dissociation at all. Since the essence of it is distinguishing fewer or more qualitative characteristics of people, it is better termed differentiation-similarization of a population. Mathematically, this concept is derived by substituting Eq. 14, Ch. III, defining differentiation into Eq. 4b,

Ch. IV, defining plurels as a product of attributes and a population. Hence the differentiation of non-group plurels (Eubank's "categories") is symbolized by

$$\underline{I}_{\Sigma, i}^0 \underline{P} = \text{differentiation of plurels} \quad (\text{Eq. 9, Ch. IV})$$

which states definitely only the number of plurels,  $|\Sigma_i$ , and leaves indefinite (as denoted by the underlining symbol) the number of persons and the nature of the characteristic of each separate plurel.

The symbols in Eq. 8, Ch. IV then, will denote sociation only when the plurels are explicitly written as, or implicitly understood to be, groups. This relationship of sociation is fundamental in Sociology, but, as its dynamic aspect, the process of sociating, is even more important, fuller discussion is postponed until the time sector is discussed.

### *E. Value Plurels, $P_v$*

In classifying indicators, it was pointed out that among psychological indicators perhaps the most socially significant basis of classification was by human evaluation into positive and negative values, i.e., desires and aversions. When these indicators characterize and define plurels, the plurels may be termed evaluated plurels or simply value plurels. Thus health is a positive value to most human beings (excepting perverts, ascetics, or psychotics). Normal people desire it and tend to behave in such a way as to avoid sickness and death (subject to the limitations of their knowledge). Sickness, crippledom, insanity, are generally negatively valued. The sick, the crippled, the insane are then negative value plurels.

$$I_v P = P_v = \text{value plurels} \quad (\text{Eq. 10a, Ch. IV})$$

$$I_{\neg v} P = P_{\neg v} = \text{negative value plurels} \quad (\text{Eq. 10b, Ch. IV})$$

$$I_{+v} P = P_{+v} = \text{positive value plurels} \quad (\text{Eq. 10c, Ch. IV})$$

The  $v$  script is simply a subvariety of the indicator script,  $i$ .

Each of the ten plurels in the segmental classification above represents a body of positive societal values. A dearth of these values, their opposites, or decrease of them, are negatively valued. Each of these ten plurels has its corresponding negative value

plurels, or social problem plurels as they are more frequently termed. Economic plurels have their problem plurels of the poor and the unemployed; educational plurels have their problem plurels of the illiterate, the ignorant, and the superstitious; political plurels have their problem extremes in the criminal within a country and the war-makers internationally; and so on down the list.

These human evaluations are subject to objective and often quantitative observation. Value plurels may then be distributed according to degrees of the value. The quantic formula for distributions is  $I^{+1}P^{+1}$ , which is the topic of Chapter V, so that here only the existence of qualitative value plurels and the notation for them will be pointed out. For examples of negative value plurels in this chapter see S. 2, 3, 4, and 9. Since, as explained later, the value script involves the subjective judgment of the analyst, *unless the evaluation has been experimentally determined*, it is not written as a v, but in the more objective notation of p, q, etc.

### III. THE POPULATIONAL CLASS-INTERVAL SCRIPT, $\text{p}|$

The multiple class-interval script in the population sector ( $\text{p}|$ ) states the number of quantitatively equal classes in which the population is expressed. The singular class-interval script ( $\text{p}'|$ ) defines a particular class-interval and, if underlined to denote indefiniteness, it denotes *any one* particular, i.e., typical, class-interval. In this sense it can assert the unit.

The unit of population is either a person, or a number of persons, a plurel. The P means the number of persons and requires no class-interval descript if that is the unit. If the unit is a plurel it must be so specified by a descript. Thus in S. 11, Ch. IV (the distribution of polygynous families in a Syrian village compared with a city), the unit of frequency is one family and is so stated by the class-interval script.<sup>9</sup> \*

### IV. THE PERSON SCRIPT, $\text{p}|$

The case, or point, script in the population sector denotes persons and is conveniently referred to as the "person script."

\* For Eqs. 11a-b, Ch. IV, see notes at end of the chapter.

It has the notational function of denoting an aggregation, or list, of persons, an array of a matrix.<sup>10</sup> \*

The person script is especially useful in specifying the structure of a group. Groups have leaders, often in an elaborate hierarchy, as in officers of an army. Groups have specialists who carry on the functions of the group in behalf of their clienteles, such as newspaper reporters and editors in behalf of newspaper readers, lawyers on behalf of litigants, doctors on behalf of the sick, etc. Some examples of the use of the person (and other) descripts in describing the personnel elements in the structure of a group may be suggestive for further uses.

Eubank cites Aristotle's structural classification of political groups (Ref. 25, p. 152). This classification on the basis of the number of rulers and an evaluation of them may be expressed in S-formulae, as follows:

Number of rulers	$\mathfrak{P}$	Positively valued indicators	$\mathfrak{P}_{+v}$	Negatively valued indicators	$\mathfrak{P}_{-v}$
					(Eqs. 13a, b, c, Ch. IV)
One	$\mathfrak{P} = 1$	Monarchy	$\mathfrak{P}_{+v}$	Despotism or Tyranny	$\mathfrak{P}_{-v}$
					(Eqs. 13d, e, f, Ch. IV)
Few	$\mathfrak{P} < .5P$	Aristocracy	$\mathfrak{P} < .5P_{+v}$	Oligarchy	$\mathfrak{P} < .5P_{-v}$
					(Eqs. 13g, h, i, Ch. IV)
Many or Most	$\mathfrak{P} > .5P$	Democracy	$\mathfrak{P} > .5P_{+v}$	"Mobocracy"	$\mathfrak{P} > .5P_{-v}$
					(Eq. 13j, k, l, Ch. IV)

Where  $P$  = the number of persons in the population (Eq. 13m, Ch. IV)  
and  $\mathfrak{P}$  = the aggregation, or list, of rulers (Eq. 13n, Ch. IV)

A second example of the adaptability of the standardized notation of S-theory is the structure of kith and kin in the family.

Let  $h$  = husbands                       $w$  = wives                       $c$  = children  
 $r$  = other relatives who may be specified with separate  
symbols as elaborately as required in a given study

Then using the standard rules for scripts and operational symbols of S-theory, the following types of families of interest to the

\* For Eqs. 12a-h, Ch. IV, see notes at end of chapter.

anthropologist and to the social worker may be expressed in formulae, thus

- a. Monogamous =  $h'+w'+cP$  (Eqs. 14a-q, Ch. IV)  
 b. Polygynous =  $h'+w+cP$   
 c. Polyandrous =  $h+w'+cP$   
 d. Group marriage =  $h+w+cP$   
 e. Partiarchal =  $h'+w+c:(w+c)P$   
 f. Patrilineal =  $h':cP$   
 g. Matrilineal =  $w':cP$   
 h. Endogamous =  $hP, :wP,$   
 i. Exogamous =  $h'P, :w'P,,$   
 j. Widow and children =  $w'+cP$   
 k. Orphans =  $cP$   
 l. Childless couple =  $h'+w'P$   
 m. "Only-child" family =  $h'+w'+c'P$   
 n. Second wife and step-children =  $h'+(w'+c'+c'')P$   
 o. Step siblings =  $c'+c''P$   
 p. Relatives =  $rP$   
 q. Divorced couple =  $h'-w'P$   
 (i.e., husband and wife subtracted) etc.

The previous example develops what is merely a symbolic notation for family structures. In itself it gives no insight into family relationships. As is the case with all symbols, their function is more to make objective and communicable the product of the intuition, the insight, or the ingenuity of the investigator. It is useful in that a clear, communicable statement of the known sharpens the unknown and the direction of research.

This function of symbols to give clear-cut expression to findings may be pointed out in the following example of a study of the probability of divorce in the United States. (Ref. 9.) For the period and the plurel studied ( $t|_p$ ) the probability of divorce of a married couple with one child was 8%, and was roughly cut in half by each subsequent child. This principle can be stated in the equation:

$$\%P = .08(.5^{c-1}) \pm \quad (\text{Eq. 15, Ch. IV})$$

where  $\%P$  is the percent of the married that are divorced, and  $c$  is the number of children as above. Thus three children would cut

the probabilities of divorce to 2%, and four children would cut it to 1%. Within the limits of the approximateness denoted by the plus-and-minus sign in the situation (and more accurately measurable by a sigma), this equation may be used for prediction. An actuarial could refine this formula and base insurance upon it. The equation gives no insight into the causes of divorce or its consequences. It simply crystallizes in a form usable for rough prediction (which is a major function of science) a net effect of whatever causes may have been operating. The pioneering, qualitative, insight-guided research into societal phenomena must go first and yield data which statistical technics and systematic symbolism, such as this S-theory, then take over to objectify, verify, and put to standardized use. Qualitative research must always precede quantification. Quantitative research then refines and verifies qualitative findings.

(For fuller discussion of creative vs. confirmatory functions of symbolism and statistics see such studies as in Refs. 31, 45, and 59.)

For a fourth example of the use of population scripts, consider the expression of findings, or principles, emerging from a study of the effect of foster homes on the IQ's of adopted children (Ref. 26). Using cross-scripts and anticipating the notation of Chapter X dealing with temporal change:

let $\mathbb{P}$   = adopted children	$\mathbb{P}', \mathbb{P}''$   = siblings
I = IQ	$t'$   = a specified age
$t$   = periods of time	$t$ I = being a change of IQ in time
	$+t$ I = a gain $-t$ I = a loss
+ and - in scripts denoting positive and negative deviations from the mean, i.e., above and below normal respectively	

The findings were:

- a.  $+t$ I, i.e., children in foster homes gained in IQ  
(Eqs. 16a-e, Ch. IV)
- b.  $\mathbb{P}I_{+p} > \mathbb{P}I_{-p}$ , i.e., children in superior foster homes gained more in IQ than children in inferior foster homes

- c.  ${}^p>{}^tI < {}^p<{}^tI$ , i.e., children who were over the average age of adoption gained less in IQ than under-age children
- d.  ${}^pI_p \cdot {}^{p''}I_p = .5$ , i.e., siblings in the same home showed a correlation of .5 in their IQ's
- e.  ${}^pI_p \cdot {}^{p''}I_q = .25$ , i.e., siblings in different homes showed a correlation of .25 in their IQ's

Instead of writing equations b and c in ordinal terms of greater or less than ( $>$ ,  $<$ ), they could be stated more exactly in cardinal terms as an equality with the numerical amount of the difference brought in to equalize the equation and with the qualifying "plus or minus the standard error of the difference" ( $\pm\sigma_d$ ).

For a fifth, neat, and simple example of populational descriptis, note the usual classification of sociological observation by scope into case studies, surveys of samples, and censuses of a whole population. These mean observing one person, a class-interval of more than one person and less than the whole population, and a specified whole population, which are denoted by each of the three descriptis, respectively, in the formula

$${}^{'P}, = \text{a case, a sample, and a whole population} \\ (\text{Eq. 17, Ch. IV})$$

#### V. THE GEOMETRY OF THE POPULATION SECTOR, $\overline{{}^2P_p}$

In interpreting the population sector and its dimensions in geometric terms, there is little to be added to the general exposition in the preceding chapter.

A population, presented in an S-situation as an aggregation of plurels, is representable as a sheaf of vectors, each plurel being one vector. See S. 35, Ch. II for a diagram of such a sheaf of vectors in n-space. In a hierarchy of plurels the lowest degree (i.e., smallest plurels written farthest to the right in the plurel script) are vectors. The next higher degree is a sheaf, the next higher degree is a sheaf of sheaves, etc. The successive degrees of sheaves from the largest down may also be thought of as successive subsectors dividing up the population sector. A population that is presented in a given S-situation as a sum of plurels is the sum of the vectors, i.e., the resultant vectors.

The class-interval script is represented by a line-sect or scalar

unit of length along a population vector. A person is the smallest unit of length on a population vector, i.e., the midpoint of the unit length representing a person. The square of a population,  $P^2$ , as in the interrelation matrix, may be geometrically represented by two vectors.

In brief then, the vectorial formula for the sector is:

$$\left. \begin{array}{l} \text{number of} \\ \text{points (persons)} \end{array} \right\} = \frac{\text{---}}{P \quad P} = \text{exponent}$$

$$P = \text{number of units of length (persons)}$$

$$\left. \begin{array}{l} \text{number of} \\ \text{line-sects} \end{array} \right\} = P \quad P = \text{number of vectors (plurels)}$$

$$= \text{the vectorial populational sectoral formula}$$

(Eq. 18, Ch. IV)

The number of dimensions in the population sector of a defined situation is given by the sum of the plurel scripts.

$$|_{p+q+\dots z} = \left| \begin{array}{c} z \\ z_p \\ 1 \end{array} \right|, \text{ or simply } \Sigma s = \text{the number of populational dimensions}$$

(Eq. 19, Ch. IV)

Thus, for example S. 2, Ch. IV has six dimensions, as it has six plurels or population components. S. 5, Ch. IV has fifteen dimensions, as there are fifteen small circles representing the lowest degree of plurels. S. 8, Ch. IV has ten dimensions; S. 10, Ch. IV has eighteen dimensions; S. 11, Ch. IV has eight dimensions (a rural and an urban sheaf, each of four vectors representing the four polygynous plurels of 1, 2, 3, or 4 wives).

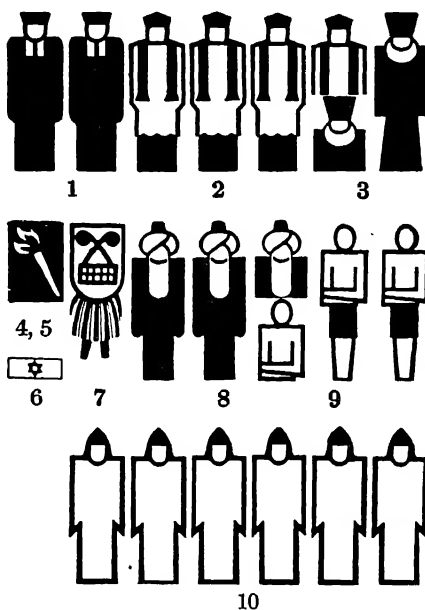
## VI. *S-SITUATIONS*

A sample of twelve situations is presented here to represent the data from which the formulae of this chapter were induced and to illustrate these formulae as applied to quantitatively recorded situations.

## S. 1

## RELIGIOUS GROUPINGS OF THE WORLD

Each whole figure represents 100 million  
persons approximately



1. Protestant; 2. Catholic; 3. Orthodox; 4. Free Thinkers;  
5. Atheists; 6. Jews; 7. Primitive Beliefs; 8. Moslems;  
9. Hindus; 10. East-Asiatic Religions.

Ref.: *Mengenbilder und Katogramme*, Gesellschafts und Wirtschaftsmuseum in Wien.

Descriptive formula:  $S_1 = P_p$

Quantic number = 0;0;0;1

Legend:

$S_1$  = The situation

classified into

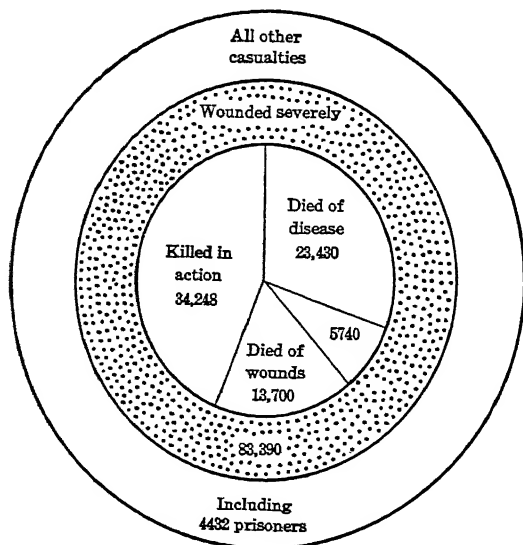
records

$|_p$  = 9 religious plurels

P = the population of the world

## S. 2

## AMERICAN CASUALTIES IN THE WAR



Final revised figures of American army casualties give a total of 302,612, including 77,118 deaths. The diagram above shows the nature of these casualties. The area of the outside circle represents the entire 302,612. The inner circle represents the 77,118 deaths, and it is sub-divided to indicate the proportion killed in action, died of wounds, and died of disease, with a fourth small slice for the deaths by accident, drowning, suicide (272), murder or homicide (154), unclassified or unknown causes, and the ten executed by sentence of general court martial. The speckled ring surrounding this circle represents the 83,390 cases severely wounded; while the large white ring at the circumference of the circle represents all the other casualties—the 91,189 slightly wounded, the 46,480 “wounded, degree undetermined,” the 4,432 prisoners, of whom 4,270 have been repatriated, and the three men missing in action.

Ref.: *Survey*, Vol. XLIII, No. 17, Feb. 21, 1920, p. 617.

Descriptive formula:  $S_2 = P_p : a$

Legend:

$S_2$  = The situation

records

$P$  = The U. S. World War casualties  
in

Quantic number = 0;0;0;1

$|_p = 3$  classes

{	deaths
	wounded
	severely
	other

with

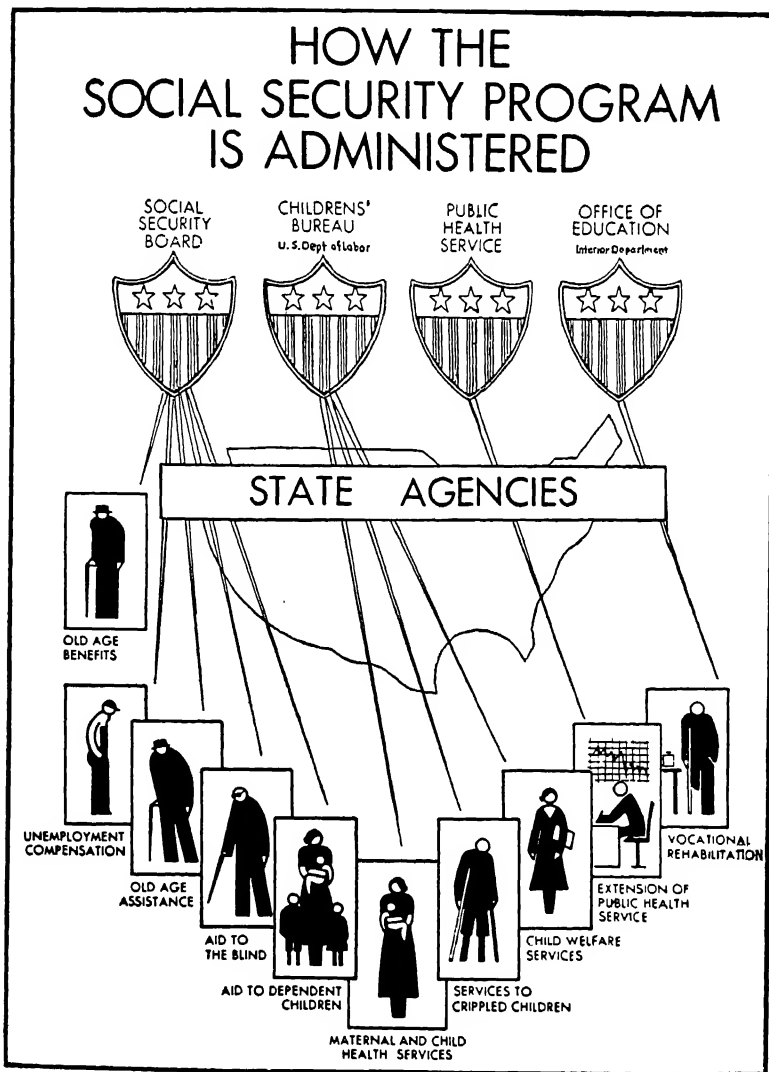
$|_a = 4$  subclasses of deaths

*Comment on notation:*

1. The irregularity of classification is illustrated in this situation in that only one of the three main classes is subclassified. The value of  $|_q$  corresponding to the two other plurels  $|_{p''}$  and  $|_{p'''}$  is zero, i.e.,  $(|_{p''}, |_{p'''}): |_q = 0$ .

2. Note that, as usual, the population is an implicit product of the *three* attributes, "American" and "casualties" and "of the World War," with a pure population, i.e.,  $I_1^0 I_2^0 I_3^0 P_0^{+1} = P_7^{+1}$ .

S. 3



Ref.: *How the Social Security Program Is Administered*, Pictorial Statistics, Inc., Social Security Board, Washington, D. C.

Descriptive formula:  $S_1 = \underline{P}_p : a$

Quantic number = 0;0;0;1

Legend:

$S_1$  = The situation : = each having corresponding  
records  $|_a$  = dependent plurels.

$\underline{P}_p$  = 4 government agency plurels  
(population indefinite)

Comment on notation:

Usually with plurels, the colon connotes subclassification in the sense of breaking a population up into smaller plurels. The present situation in which the colon means subclassification in the sense of jurisdiction or clientele is unusual.

## S. 4

### WASTES IN CONSUMPTION

	Estimated Man-Power	
	Total	Wasted
1. The military establishment . . . . .	1,500,000	1,000,000
2. The opium and cocaine traffic . . . . .	unknown total	
3. The drug traffic . . . . .	400,000	100,000
4. Distilled spirits . . . . .	unknown total	100,000
5. Prostitution . . . . .	250,000	150,000
6. Crime—criminals . . . . .	320,000	200,000
—watchers of criminals . . . . .	400,000	200,000
7. Adulteration . . . . .	unknown total	
8. Speculation and gambling . . . . .	unknown total	
9. Quackery . . . . .	unknown total	
10. Super luxuries and fashions . . . . .		6,000,000
11. Commercialized recreation . . . . .	unknown total	
12. The overhead services . . . . .	unknown total	
13. Advertising . . . . .	600,000	250,000
Total minimum wasted man-power . . . . .		8,000,000

Ref.: Chase, Stuart, *The Tragedy of Waste*, The Macmillan Company, 1925, p. 106.

Descriptive formula:  $S_4 = P_p : a$

Quantic number = 0;0;0;1

Legend:

$S_4$  = The situation  
records

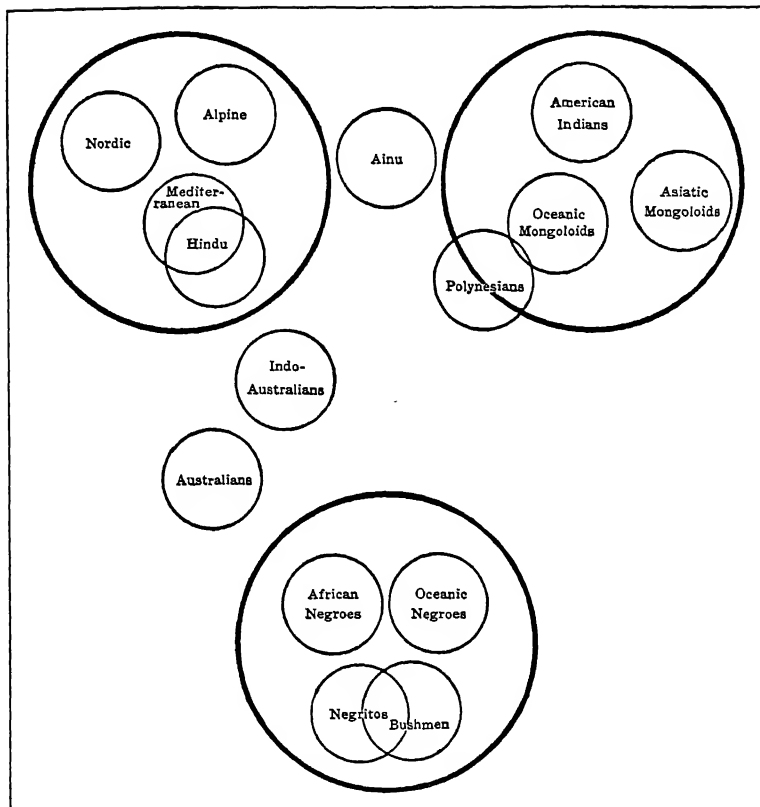
$|_p$  = 2 plurels  $\left\{ \begin{array}{l} \text{total} \\ \text{and} \\ \text{wasted} \end{array} \right.$

$P$  = man power  
in

in each of

$|_a$  = 13 occupational plurels

## S. 5



Ref.: Kroeber, A. L., *Anthropology*, Harcourt, Brace and Company, 1923.

Descriptive formula:  $S_5 = P_p : q$

Legend:

$S_5$  = The situation  
records

P = the population of the world  
classified into

Quantic number = 0;0;0;1

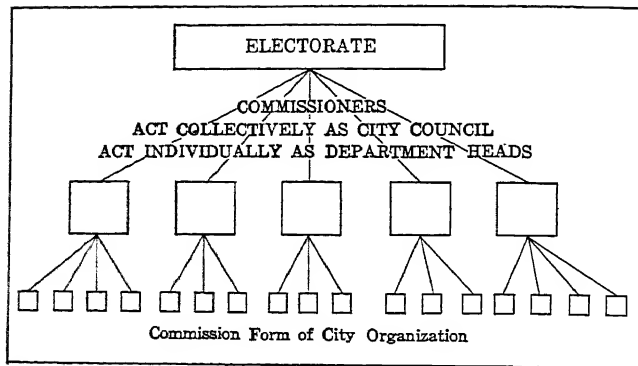
$|_p$  = 4 continental  
plurels

{ Eurasia,  
America,  
Africa, in-  
between

and subclassified into

$|_q$  = the chief races, 15 in all.

## S. 6



Ref.: Pfiffner, John M., *Public Administration*, The Ronald Press Company, 1935, p. 36.

Descriptive formula:  $S_6 = \underline{P}_p : q : r$

Legend:

$S_6$  = The situation

records

$: :$  = a hierarchy of plurals

with

$\underline{P}_p$  = the electorate at the top

Quantic number = 0;0;0;1

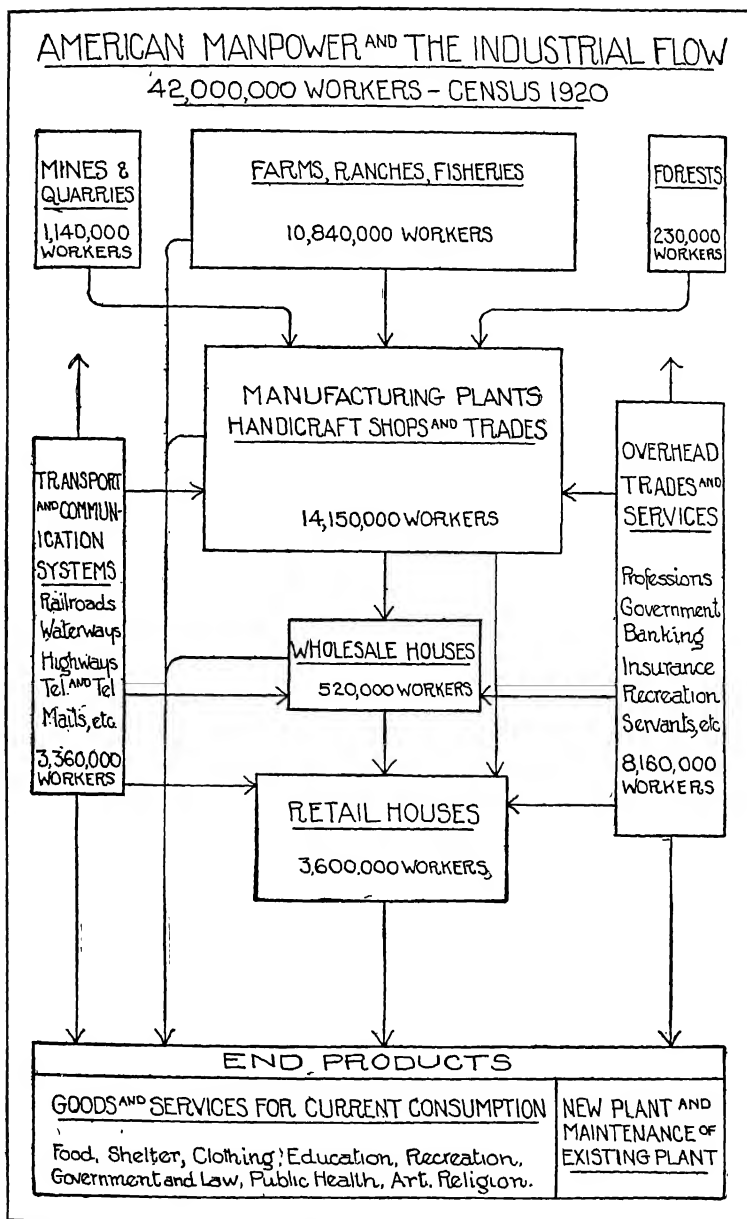
with

$|_q$  = the Commission and its 5  
Departments subordinate

and

$|_r$  = sub-Departments under them

## S. 7



*Descriptive formula:*  $S_i = P_p \cdot q \cdot r$

*Quantic number* = 0;0;0;1

*Legend:*

$S_i$  = The situation

subdivided into

records

$P$  = the U. S. population

$|_q = 6$  subtypes

divided into

{ mining, farms,  
forests, trans-  
port, mfg. and  
selling, overhead

and

$|_p = 3$  economic  
plurels

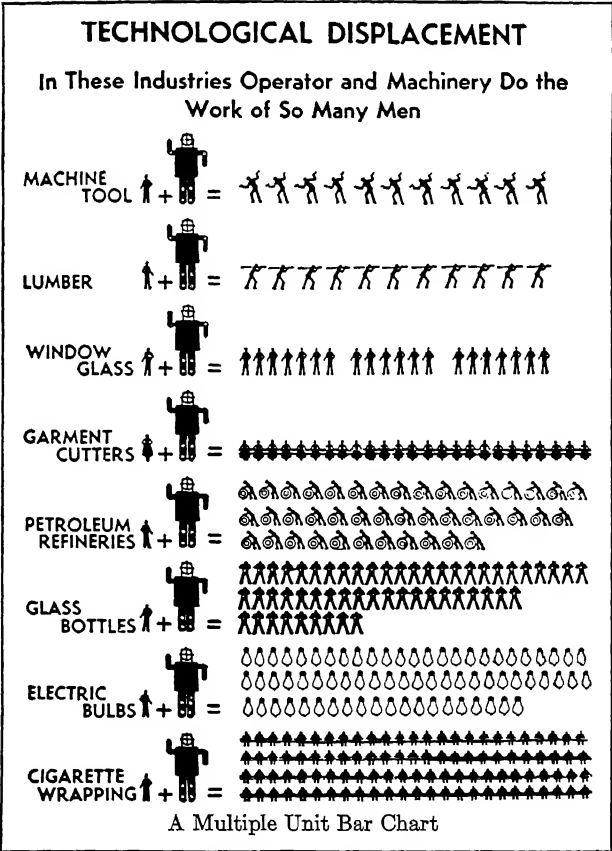
{ extractive  
producers  
mfg., dis-  
tributing  
and service  
producers,  
consumers

$|_r = 3$  further subtypes—mfg.,  
wholesalers, retailers.

*Comment on notation:*

The date should be included in the formula, but until the time sector and its scripts are discussed further on, it is omitted here for simplicity of exposition in developing one sector at a time.

S. 8



Ref.: Arkin, H. and Colton, R. R., *Graphs: How to Make and Use Them*. Reprinted by permission of Harper & Brothers, 1936, p. 104.

Descriptive formula:  $S_8 = "P_p$  Quantic number = 0;0;0;1

Legend:

$S_8$  = The situation in each of  
records  $|_p$  = 8 industrial plures  
 $P$  = the number of displaced work- compared with  
ers  $"|$  = the one worker replacing them

Comment:

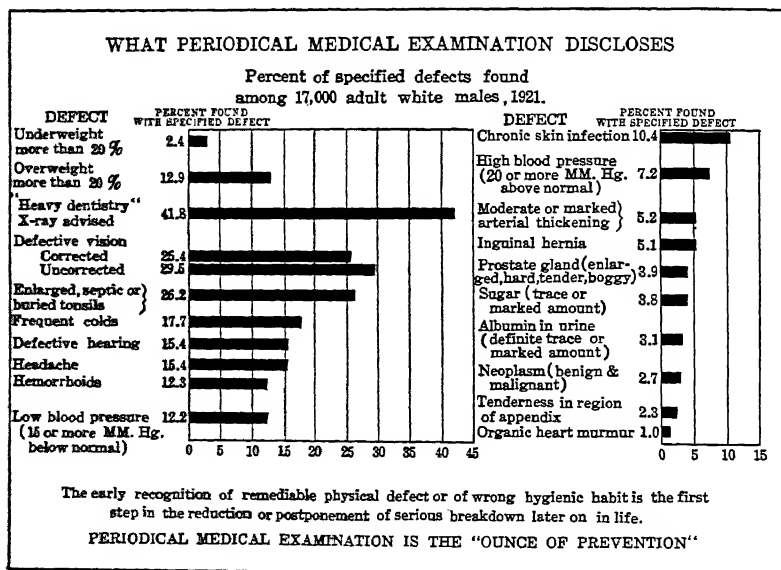
The simple equation above describes the message of the graph, which is to show the number of displaced workers by industries. The situation's implications may be better analyzed and written in a "calculative" equation as follows:

$$| (P \times I^0) I = P^M I |_i$$

where 'P' is the single unidentified worker with the machine,  $I^0$ , of each of 8 kinds,  $|_i$ . The equation requires a common unit which is implied in the term,

"the work," in the caption, meaning the amount produced. Let this amount produced by the machine and its one worker (in units appropriate to each industry) be the indicant I, and the former mean or per capita production be denoted by  $\mathbb{M}$ . Then the left-most member of the equation represents the total production of the machine and its one worker which equals the right member representing the total production of the P displaced workers without the machine. Whether such equations are sterile statements of these relationships, or whether they can be fruitfully manipulated so as to increase precision in thinking, or to reformulate problems for further investigations, or to suggest generalizations, or to reveal further unsuspected relations, is a methodological question not yet fully explored.

## S. 9



Ref.: Davis, Jerome, and Barnes, Harry E., *An Introduction to Sociology*, D. C. Heath and Company, 1927, p. 757. Courtesy of Louis I. Dublin, Metropolitan Insurance Company.

Descriptive formula:  $S_9 = \%P_p$

Quantic number = 0;0;0;1

Legend:

$S_9$  = The situation  
records

in each of  
 $|_p$  = 21 defective plurals

$\%P$  = the percent of 17,000 white  
adult males

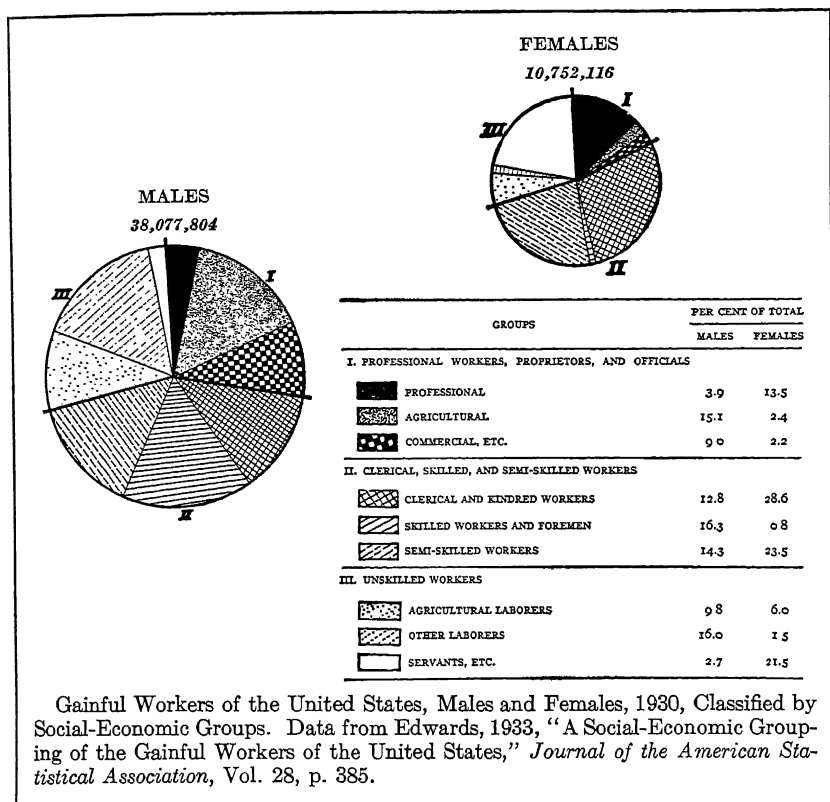
Comment:

The percentages of subnormal persons define the "minimals" (i.e., social problem plurals) in these 17 characteristics of health.

Comment on notation:

As in S. 4, Ch. IV, the date is omitted in the descriptive formula for simplicity, until the temporal notation is studied in later chapters.

## S. 10



Descriptive formula:  $S_{10} = \%P_p : q : r$

Legend:

$S_{10}$  = The situation

records

P = the U.S. employed population

divided into

$|_p$  = 2 sex plurels

and

Quantic number = 0;0;0;1

$|_q$  = 3 occupational plurels

and

$|_r$  = 3 occupational subplurels

expressed in

$\%|$  = percents

Comment on notation:

1. Note that the colon usually, in the population sector, denotes subclassification in the sense of breaking up a plural into its constituent plurels.

2. The situation is a matrix of the third degree as denoted by its 3 multiple aggregative scripts. It has 18 dimensions ( $p \times q \times r = 18$ ), the number of its qualitatively different plurels.

## DISTRIBUTION OF ANSWERS TO QUESTION (NO. 10) ABOUT POLYGYNY

The "normal" rural Arab sample (N = 100) (families identified by number)

Compared with the urban sample (N = 50) (each family indicated by two commas)

Number of simultaneous wives per husband	0	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1	02,06,07,10,13, 03,06,08,11,14, 18,26,31,34,38, 40,42,44,47,49, 51,53,57,61,63, 67,74,76,78,80, 85,88,91,96,	17,24,27,32,37, 39,41,43,45,48, 52,54,58,62,66, 69,75,77,79,82,	85,88,91,96,								
2	11 16 33 36 71 73 90										
3	15 30 35 50 72 83 99										
4	04 93 97 84 95 00 59 81										
Frequency	0	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Average number of wives per married man = 1.25. 13% of the families are polygynous in these villages (all families are nominally Moslems). All are monogamous (Christians) in the urban sample. Among these Alaouites bachelorhood, monogyny, or polygyny are chiefly determined by income. The men tend to marry as early, and as often, as they can afford to do so.

Ref.: Dodd, Stuart C., *A Controlled Experiment on Rural Hygiene in Syria*, American Press, Beirut, 1934, Table 16.

with a corresponding

*Comment on notation:*

1. The attribute symbolizing the statements of attitude is written explicitly because it is a second attribute. The first attribute, representing the qualitative characteristics defining the four plurels, has been multiplied by the "pure" population, giving an attribute-population product in which, as usual, the implicit attribute is denoted by the descripts on the population index.

## VII. NOTES

1. The operation of division can be denoted mathematically by a negative exponent:

$$P_i/P_{i++} = P_i^{-1}P_{i++}^{-1} = {}_{\infty}P \quad (\text{Eq. 1e, Ch. IV})$$

This denotes a population of persons,  $P$ , in number, expressed in units of a more inclusive population of persons,  $P_{i++}$ , in number. (See S. 9, 10, 11, and 12, Ch. IV, for examples.) A death rate expressing the ratio of the number of persons dying to the number of persons in the population; a percentage of persons possessing some characteristic in a population, are examples. The minus exponent, denoting a divisor, never exists alone in a situation. As the combination of the exponents of minus one and plus one merely changes the units in which the population is expressed, the ratio is practically the same as a simple primary index, so that the quantic digit is considered to be 1. But, if the ratio is between two *different* populations, as in a fertility ratio of babies per mother, the plus and the minus exponents are recorded by the digits 19 in the quantic number. Thus the quantic number of a situation

with a $P^0$ in it is	e;e;e;0	} nullary	(Eq. 2a, Ch. IV)
$P^{+1}$	e;e;e;1		(Eq. 2b, Ch. IV)
$P^{-1}$	e;e;e;9	} primary	(Eq. 2c, Ch. IV)
$P^{+1}$	e;e;e;19		(Eq. 2d, Ch. IV)
$P^{+2}$	e;e;e;2	} secondary	(Eq. 2e, Ch. IV)

Digits of 1, 9, and 19 in the quantic number all denote subvarieties of a primary index.

Digits of 2 and 8 in the quantic number denote subvarieties of a secondary index.

2. "Cross-scripting" consists of describing an index in one sector by means of a script from another sector. The resulting compact "Brief-S" formulae are treated in full in Chapter IX.

3. This concept of a described  $P$  seems to the author to be much the same as the concept of "the situation self," developed by Eubank (Ref. 25, p. 106) as the person with those of his characteristics that are involved in a given situation.

4. The term "degree" for the successive strata levels, or orders of a classification, is used as having fewer undesired connotations. The term "orders" is pre-empted in matrix algebra to mean the number of rows and columns in a matrix. Thus a matrix of 4 rows and 3 columns is said to be of order  $4 \times 3$ . Such a matrix has but two "degrees," i.e., the rows and the columns. If it had sagittal arrays, it would become a third-degree matrix, as, for example, one of

order  $4 \times 3 \times 6$ . "Strata" and "levels" depend upon viewing the classification in the form of the diagram above. They also have evaluative connotations of caste and social class in sociological usage.

"Degrees" are identifiable with "subsectors," but "subsectors" connote co-ordinate entities as well as superordinate-subordinate entities. The term "dimension" might be extended to include these "degrees," but then "dimension" would denote both indices and groupings of indices, vectors, and sheaves of vectors, and confusion would result. To symbolize the concept of degrees:

$$\begin{aligned} \text{Let } |_{\Sigma S} &= \text{the number of degrees in all sectors} & (\text{Eq. 5a, Ch. IV}) \\ \text{and } |_{\Sigma s} &= \text{the number of degrees in one sector} & (I') \quad (\text{Eq. 5b, Ch. IV}) \end{aligned}$$

Thus in Eq. 4c, Ch. IV above, since four letters are written in the class script,  $|_{\Sigma s} = 4$ . This number of degrees of the matrix (and strata of the hierarchy also) should not be confused with the total order of the matrix, described in the preceding chapters, §1.

5. A list of adjectives is comparable to a product of several attributes, such as "American, gainfully employed, male, unskilled, agricultural persons," in S. 10, Ch. IV, and would be symbolized in the condensed notation as:

$$P_{p,q,r,s} = \text{a product of five factors which are four attributes and the plural they characterize.} \quad (\text{Eq. 6, Ch. IV})$$

This illustrates how qualitative entities are multiplied together. This is comparable to what logicians call "the logical product" (Ref. 80). If the frequency of occurrence in a population of each attribute-adjective were known and expressed as a proportion of the frequency of its occurrence plus its non-occurrence, it would be the probability of that attribute. The product of several such probabilities is their joint probability or probability of simultaneous occurrence. The product of several attributes is such a joint probability in the incipient stage before the frequency of the characteristic, which the attribute represents, has been observed. The attribute concept thus helps to bridge the gap between verbal and literary descriptions and mathematical ones. It defines an operational stage of observation which is more precise than the verbal adjective, yet not as precise as the ordinal, cardinal, and calibrated quantities outlined in our theory of measurement in Chapter III.

6. The obverse question of when the population is implicit and when it is nul is more difficult, as there are borderline situations, and the writing or not writing of the P explicitly changes the quantic cell into which that situation is classified. The rule is to write the P explicitly if the *number* of persons or plurals is asserted (from actual observation, or as a theoretical proposition). But it is often difficult for the analyst to decide whether an enumeration or canvass is essential when some plural is referred to, or only implied in the situation as presented. For concrete cases consider the following situations in which the present author and his assistants analyzed the P either as explicit or nul, but with which another analyst might not agree:

- P explicit—S. 19, Ch. II; S. 6 and 7, Ch. VII; S. 11 and 13, Ch. 8  
 P nul —S. 4, Ch. III; S. 4, 17, and 19, Ch. VIII

A further research study might develop a scale for rating the explicitness-implicitness of the population index in borderline instances and so define the  $P$  more objectively.

7. (a) The family, or domestic group; (b) the economic, or productive group; (c) the state, or sovereign, governmental group; (d) the church, or religious group; (e) the school, or educational group; and (f) the play, or socio-recreational group; and (g) others such as: caste, class, crowd, gang, mob, party, public, and sect.

8. Cooley's we-groups and they-groups, and in-groups and out-groups generally, Miller's vertical and horizontal groups, the social classes, castes, and many other classifications which are basically defined by people's attitudes towards each other, can be more precisely defined by indicators of attitude between persons and plurels, such as social distance tests, invitations, memberships, costumes, etc. All of these involve  $P^2$  in the quantic formula and, therefore, belong in Chapter VII.

Primary or face-to-face groups vs. secondary groups involve criteria of spatial proximity and frequency of recurrence of contacts in time and will, therefore, be deferred till the  $L$  and the  $T$  components have been discussed.

The reader should realize that this separation of the phenomena of plurels generally, and of groups more narrowly, into different parts of this book is a consequence of necessities of exposition and is not involved in the quantic classification. In the quantic solid (see S. 33, Ch. II) they all occupy one contiguous region with subcells, depending on whether interrelation, spatial proximity, or dynamic processes are involved. But exposition unfortunately has to be "linear," proceeding sentence by sentence, page by page, chapter by chapter, in some orderly route through any complex of phenomena. For pedagogic and logical reasons we have chosen the route of proceeding from sector to sector, cumulatively studying situations with more and more sectors, until only at the end, with the exposition of all the four sectors completed, is the student equipped to turn around and deal with any and all situations regardless of what their combination of explicit and nul indices may be.

9. For other examples, see S. 27 and 33, Ch. VI; S. 36 and 73, Ch. X; S. 8 and 27, Ch. XI.

$\mathfrak{P}_p$  = a population of  $|_p$  plurels, each measured by subplurel  $|_q$  as units

(Eq. 11a, Ch. IV)

$\%P_p$  = a population of  $|_p$  plurels in each of which there is a percentage of plurels of the kind  $|_q$

(Eq. 11b, Ch. IV)

10. Some refinements of the person script for typical persons, and other special notational uses, are:

$P$  = a number of persons (Eq. 12a, Ch. IV)

$\mathfrak{P}P$  = an aggregation or list of persons,  $\mathfrak{P}|$  in number, an array of a matrix with a person in each cell (see S. 4, Ch. X) (Eq. 12b, Ch. IV)

$\%P$  = a population expressed in percentage units (Eq. 12c, Ch. IV)

$\mathfrak{P}^*P$  = in addition to a number of persons  $P$ , specified persons,  $\mathfrak{P}|$ , are asserted (Eq. 12d, Ch. IV)

- $^pP$  or  $'P$  = a single identified person (see S. 10 and 15, Ch. II; S. 9, Ch. IX;  
S. 23, Ch. XI; S. 24, Ch. XII) (Eq. 12e, Ch. IV)
- $'P$  = a single unidentified, i.e., typical, person (see S. 31 and 46, Ch. X)  
(Eq. 12f, Ch. IV)
- $^pP$  = in addition to a population  $P$ , certain typical persons are specified  
(see S. 24, Ch. II; S. 14, Ch. X) (Eq. 12g, Ch. IV)
- $''P$  = in addition to a population  $P$ , a certain identified person is specified  
(see S. 3 and 4, Ch. VII) (Eq. 12h, Ch. IV)

## Chapter V

### DISTRIBUTIONS, I; P

#### I. DISTRIBUTIONS OF INDICANTS

##### A. List vs. Frequency Distributions of Persons and Plurels

A plurel was defined as the product of an attribute and a pure population, i.e., a qualitatively characterized number of persons. A societal distribution may be defined as the product of an indicant (or, more generally an index) and a population, i.e., a population whose parties are characterized by quantities of some quality. As the first operational step in compiling a distribution is to observe for every person his corresponding quantity of some characteristic, the simplest type of societal distribution is a list of persons, each with the number of his I-units, of some indicant.

When the list has been compiled, the next step is to rearrange it into a frequency distribution showing the number of parties at each class-interval of the indicant.

$I : P$  = a frequency distribution of persons (Eq. 1a, Ch. V) <sup>1\*</sup>

##### B. Moments Defining the Types of Distributions

Any societal situation whose quantic formula classifies it as a distribution has properties which are best summarized by considering its dimensions as given in the form of the statistical moments. For any specified distribution of parties in some characteristic, the first question is as to the size of the population distributed. The second question is as to the representative or average amount of that characteristic in that population. The next question asks the amount of scatter of the parties on either side of that average. Just how accurately does that average represent all the individual parties in the distributed population? A fourth question is as to the lopsidedness or asymmetry of the scatter of the distribution. Are the parties more widely scattered

\* For Eqs. 1b-6, Ch. V, see notes at end of the chapter.

or dispersed on one side of the average than on the other? These four questions determine the properties of *population*, *central tendency*, *dispersion*, and *skewness* of a distribution.

Among the various indices measuring central tendencies, dispersion, and skewness, the most useful and standardized indices, derived from the successive statistical *moments*, are defined as follows:

$\Sigma I^0 P^{-1} = 1$  = zero-order moment, the unit-population, or unit-area under the distribution curve (Eq. 7, Ch. V)

$\Sigma I^{+1} P^{-1} = M$  = the first-order moment about zero, the arithmetic mean, the best single index to represent the whole distribution (Eq. 8, Ch. V)

$\Sigma {}_d I^{+2} P^{-1} = \sigma^2$  = the second-order moment about the mean ( ${}_d I$  = units of deviation from the mean =  $i - M$ ) the variance, or standard deviation squared.  $\sigma$ , sigma, is the best measure of dispersion (Eq. 9, Ch. V)

$\Sigma {}_\sigma I^{+3} P^{-1} = \sqrt{\beta_1} = Sk_1$  = the third-order moment about the mean, in standard units,  ${}_\sigma I = {}_d I / \sigma$ ,  $Sk_1$  is a measure of skewness. Pearson's  $\beta_1$  is a criterion of the shape of the distribution curve (Eq. 10, Ch. V)

$\Sigma {}_\sigma I^{+4} P^{-1} = \beta_2$  = the fourth-order moment about the mean in  $\sigma$  units, a measure of kurtosis (peakedness) and a criterion of the type (i.e., shape) of the distribution curve.<sup>2\*</sup> (Eq. 11, Ch. V)

## 1. THE DIMENSIONS OF THE MOMENTS

The "orders" of these indices are specified by the indicatory exponent. Thus, the mean is a dimension to the first power, the variance is a dimension squared,  $Sk_1$  is a dimension cubed, etc., as shown by the positive exponent,  $|^{i=3}$ . The square of a dimension is geometrically represented by a line and a normal (perpendicu-

\*For Eqs. 12a-13, Ch. V. see notes at end of the chapter.

lar) to it; the cube by combining a normal with the plane of the other two, etc. Thus, raising a dimension to a power means utilizing as many dimensions of space as there are units in the exponent. The number of dimensions in a situation is stated by the class script for the vectorial dimensions and by the exponent for any additional normals or exponential dimensions, as they might be called. The variance of an index, area, and volumes of length are the commonest examples. These can be shown to be special cases of the general scalar product of vectors. They are correlation of an index with itself in statistical terms, or the scalar product of a vector with itself in vector terms, or a dimension squared in terms of dimensions.

The formula for a distribution has two dimensions. Its positive quantic digits which are the exponents, add up to two ( $I^{+1}$ ;  $P^{+1}$ ). A two-dimensional surface is required to graph a distribution. The indices describing the distribution range from one dimension up, with decreasing usefulness for the higher dimensions. The indices of the first and second dimensions, e.g., the mean and sigma, are the most useful and the most used. In the Quantic Table, Chapter II, arrays are shown only for  $I^0$ ,  $I^{+1}$ , and  $I^{+2}$ ,  $I^{+3}$  measuring skewness,  $I^{+4}$  measuring kurtosis and higher orders are still rare in the sociological literature and are therefore omitted in the table, though place for them is provided and they can be inserted whenever needed.

### *C. Averages and Norms*

The arithmetic mean includes all data expressed in units per capita or per plurel. It includes all percentages since these are but means of all-or-none indicants. (See Eq. 17, Ch. III.) It is the index representing the amount of a characteristic of a plurel, just as an indicant represents the amount of a characteristic of a person.<sup>4</sup> In all comparative work in Anthropology and Sociology the characteristics of plurels may be compared by noting the difference, or sometimes the ratio, of their means. For value indicators the mean is one norm (an ethical standard).

Persons near the mean in some characteristic are "normal." Extreme deviations from the mean define the "abnormal." The mean of value indicators is thus seen as a most important sociological measure.

*D. Dispersion*

## 1. RELATION TO PROBABILITY

The standard deviation, also called sigma ( $\sigma$ ), is a point on the indicant scale, °I, which is conventionally standardized as the best general index of dispersion, or scatter, of the parties on both sides of the mean. It has many sociological uses. In the normal distribution it is that deviation from the mean which includes one third of the population. Thus, two thirds of the population are characterized by an amount of the indicant which is within plus or minus one sigma. In S-theory  $\sigma$  is taken as defining the limits of the "normal" part of a population in a distribution.

$\pm\sigma+^M I : P = .68P$  = The definition of "normal" persons in a normal distribution (Eq. 14, Ch. V)

(See S. 44, Ch. X, for a table of such norms for a set of intelligence indicants, and S. 12, Ch. V, for a graphical meaning of sigma as a deviation from the mean.)

For any person taken at random from a normal distribution this defines the probability of his being a normal person, i.e., the chances in a hundred of his having an amount of the distributing characteristic which is between one sigma below the mean and one sigma above the mean. More refined probabilities of a person reaching, exceeding, or being between specified points of the indicant are given by the table of the normal probability distribution.<sup>5 \*</sup>

## 2. RELATION TO RELIABILITY

A major use of this interpretation of sigma in probability terms is in expressing the reliability of any sociological index. Such indices are calculated from samples of an actually or conceivably more inclusive population, called "the universe sampled." If repeated samples of a given size, P, are taken, their indices will show a fluctuation which is called "sampling error." The extent of these fluctuations is measured by the standard deviation,  $\sigma$ , of the distribution of that index in a population of very many samples. This  $\sigma$  is called the standard error. Every statistical index may have a standard error calculable from a formula.

\* For Eq. 15, Ch. V, see notes at end of the chapter.

Scientific "good manners" require that, when indices are stated in any study, the standard error of those indices be reported as this measures the degree of unreliability of those indices as far as sampling error is concerned. For given amounts of standard error the corresponding probabilities of their occurrence may then be read from the table of the probability integral. Many sociological studies would be more accurate if scientific etiquette in this respect were observed.<sup>6</sup>\*

### 3. RELATION TO COMPARABILITY OF UNITS

A corollary of the standard deviation being a unit of the indicant, which is convertible into units of probability, is that the standard deviation thus becomes a common unit by which different kinds of indicants can be compared, averaged, correlated, etc. Thus, different characteristics of individuals or of plurels, which are expressed in a great diversity of units, can be re-expressed in sigma units of the same or of a standard population. This makes possible such comparisons, or predictions, as:

1. John is more intelligent than he is musical, i.e., his IQ in sigma units is greater than his Seashore test score in sigma units.
2. City A has a Department of Public Health which is more efficient by  $.7\sigma$  than its Department of Education, i.e., each department is rated by one of the existing elaborate and well validated scales, and its score converted into sigma units in a distribution of many cities.
3. The interest of Plurel A in foreign missions exceeds its interest in local government by  $.2\sigma$ , but is less than its interest in reading by  $.3\sigma$ , i.e., taking as indicators of these characteristics dollars per capita given to missions annually, percent of voters voting, and number of books per capita borrowed from free public libraries, and finding the distribution of each indicant in some common population of plurels, the indicants of Plurel A expressed in sigma units can be quantitatively compared.
4. Hottentots excel over Zulus by 15 I-units in characteristic x, but this observed difference is not significant, i.e., it is only

\* For Eqs. 16a-d, see notes at end of the chapter.

- .67 times its standard error <sup>7</sup> \* and so might occur by fluctuations of sampling 50 times in 100.
5. In a population comparable to the U.S. army draft of 1917, men with an intelligence rating one sigma above the mean will have a most probable fertility of  $x$  children. (See S. 35, Ch. VI.)
  6. Under conditions as in the United States from 1860 to 1910, the notables per 100 persons born in a State may be predicted from a knowledge of, (a) the population density, (b) elementary education, (c) capital per square mile, and (d) coolness of climate, with 96% of determinateness, i.e., the multiple correlation coefficient between notables and the four environmental conditions is .98.<sup>8</sup>

These and thousands of similar comparisons, measurements, and predictions have become possible through the invention of a standard unit of dispersion such as sigma is. There are other units of dispersion, such as the average deviation, quartiles, percentiles,<sup>9</sup> the probable error, etc., but the standard deviation, as its name implies, is the most general or standard unit. It is a probability unit when the shape of its distribution is known. In this sense it is a unit of unusualness of occurrence, of difficulty (if the indicant measures an achievement involving effort), or of *relative* excellence (for value indicants).

The more readily the student of sociology can think of societal characteristics in terms of sigma units of a defined population, the greater will be his facility in relating and predicting such characteristics. Ordinary indicants do not have a predictive connotation. Sigma indicants, because they are convertible into probabilities when their distribution is known, have the connotation of prediction. Prediction of the future is both a major human interest and a function of science. Biologically viewed, the ability to predict means more adequate adjustment of man to his environment.

#### 4. RELATION TO MAXIMAL DISPERSION (MONOPOLY) $\sigma_v = \Sigma VP^{-.5}$

Very often the statement of the amount of the dispersion of some characteristic seems meaningless to the student for lack of associations in that field, i.e., for lack of a standard by which to

\* For Eqs. 17-18b, Ch. V, see notes at end of chapter.

judge whether a sigma of  $x$  units is large or small for that kind of characteristic. To remedy this lack, the percentage of a sigma to the maximum sigma is here proposed. Dispersion becomes maximal in a monopoly when one party has all of the indicant, and all the other parties have zero amount of the indicant. Dispersion thus varies between the limits of equality ( $\sigma_v = 0$ ) and monopoly.<sup>10</sup>\*

Thus in S. 17, Ch. II, to say that the dispersion of life insurance among the ten leading nations is 569 million yen is not illuminating. But to state that it is 67% of a complete monopoly, with the United States leading, is more significant.

Again, this ratio enables comparison of diverse situations measured in diverse units. It becomes possible to compare the dispersion of life insurance (originally expressed in yen units) with the dispersion of intelligence (observed in score units) by showing that one is  $x\%$  nearer its maximum than the other to its maximum. This monopoly index measures the degree of dispersion of any characteristic on a percentage scale where 0% represents equality of distribution ( $\sigma = 0$ ) and 100% represents maximal inequality of distribution

$$100 \sigma_v / \text{max. } \sigma_v = (\text{mon-I}) = \text{index of the percentage of monopoly} \quad (\text{Eq. 20, Ch. V})$$

## 5. RELATION TO "DISSIMILARIZING" AND TO "VARIATION"

Dispersion, measured by sigma as the best general index of deviations from the central tendency, has been shown above to connote: (a) probability of occurrence which is useful in prediction, (b) reliability of observation which is essential in methodology, and (c) comparability of measures which is also essential in methodology.

One further property of dispersion needs to be noted to distinguish this concept from "dissimilarizing" and from "variation." Dispersion as defined here denotes quantitative variation, i.e., the deviation of the distributed units from the mean of the distributing characteristic by some amount, i.e., some number of those distributing units. Dissimilarity is defined (see Ch. III) as qualitative variation, i.e., the number of different kinds of

\* For Eqs. 19a-b, Ch. V, see notes at end of the chapter.

characteristics existing in a situation. Geometrically stated, dispersion states the length of one vector, while dissimilarity states the number of vectors, i.e., the number of different directions or dimensions of societal space that is involved in any defined S-situation. The term "variation" is used here to denote either or both types, namely, dispersion and/or dissimilarity.

$$|_i = \text{"dissimilarity"} \quad (\text{Eq. 21a, Ch. V})$$

$$^{\circ}I = \text{"dispersion"} \quad (\text{Eq. 21b, Ch. V})$$

$$^{\circ}I_i = \text{"variation"} \quad (\text{Eq. 21c, Ch. V})^{11}$$

In applying these definitions we would term as "variation" what North terms "social differentiation" (Ref. 50). He distinguishes four types, namely, by rank, occupation or function, culture, and interest. Rank is usually an ordinal indicant (although occasionally it is a cardinal, as in economic rank measured by monetary income) and is, therefore, a matter of dispersion on a quantitative scale, whereas function is mostly qualitative and is, therefore, our "dissimilarity." His "culture" and "interest" may have quantitative as well as qualitative aspects, and hence involve both our "dispersion" and our "dissimilarity." Hence, Eq. 21c, Ch. V, includes all four of North's types of societal variation.

There are numerous sociological terms for dispersion in particular forms: "stratification," "superordination and subordination," "social classes," "inequality," "vertical social distance," "rank," "income class," "educational and cultural level," "hierarchy." These may be readily analyzed into qualitative characteristics showing them to contain an element of our dissimilarity,  $|_i$ , and a quantitative characteristic measurable on a rating scale, if no more precise units are available. The degree of "superordination," "levels," "ranks," etc., of any person, or plurel, should then become more exactly expressible as a certain number of sigma units above or below the mean of some defined population which yields the norm.

### *E. Skewness and Kurtosis*

The two next higher order indices of a distribution are Pearson's  $\beta_1$  and  $\beta_2$ , defined by Eqs. 10 and 11, Ch. V. The square root of  $\beta_1$  is a measure of the skewness (Sk) of the distribution

curve. Skewness is negative when the negative tail is drawn out, as in the distribution of marital satisfaction (S. 1, Ch. V). Skewness is positive when the positive tail is skewed out to the right, as in the usual distribution of income (see S. 20, Ch. XII). Skewness may connote that the moments of the distribution are not in balance, that there may be instability in the factors creating the data, or that a change may be going on in the internal structure of the situation. This hypothesis of the possible meaning of skewness would lead, in situations where the hypothesis may be verified, to its becoming a useful tool for diagnosing evolutionary trend and probable stability of distributed phenomena.

Skewness may help to differentiate imperfectly defined characteristics. Thus, psychiatrists admit a considerable percentage of uncertainty in the diagnoses of schizophrenia and the manic-depressive psychoses. These psychoses have been found to show skewed as against non-skewed distributions respectively in their rates of incidence in Chicago when studied by ecological zones. (Ref. 21.) Thus, the dimensional analysis of phenomena may help to classify and consequently understand and eventually control those phenomena more adequately.

### 1. THE J-CURVE HYPOTHESIS OF CONFORMING BEHAVIOR

Skewness may also be due to one or a few causal factors of dominant importance being combined with a larger number of uncorrelated lesser factors. Thus, in Allport's J-curve hypothesis of conforming behavior (Ref. 1) this type of skewing was observed wherever persons were chiefly, but not entirely, controlled by a single stimulus or motive, such as stopping a car entirely at a traffic stop sign, or arriving on time at work in the morning.

As an example of kurtosis-type, Allport's hypothesis that conforming behavior, when distributed on a non-telic continuum, tends to show leptokurtic (i.e., high peaked) curves, has been tested by Dudycha (Ref. 20). The latter collected additional data of punctuality distributions of the time of arrival of students at classes, chapel, breakfast, etc., and tested both his and Allport's distributions by more exact statistical technics than the previous technic of inspecting graphs, and found them to be prevalingly mesokurtic (i.e., between peaked and flat-topped).<sup>12</sup> \*

\* For Eqs. 22a-23, Ch. V, see notes at end of the chapter.

Such tests are technical and unfortunately, to many sociologists, meaningless. But they are essential for defining the properties of phenomena, and classifying them into categories whose causal factors can then be more objectively isolated and better predicted when the type of distribution is determined.

### *F. The Normal Curve*

An important special case among the types<sup>13 \*</sup> of distribution curves is the normal probability distribution. To suggest the sociological importance of this normal distribution, two hypotheses are offered which interpret its meaning.

#### 1. THE MULTIPLEX-ELEMENTS HYPOTHESIS

The normal curve will result from a very large number of uncorrelated, more or less equal, elements. Gauss used this concept, stated mathematically, in deriving the equation of the normal curve (as Bravais also did in deriving the equation of the normal correlation surface in 1846). The normal curve is the limiting case of the point-binomial distribution, which can be typified by pitching groups of pennies. The frequent occurrence of normal distributions in biological and psychological characteristics is commonly accepted as evidence for the hypothesis that such characteristics are caused by a multiplicity of small, equal, independent elements. Let us designate this hypothesis of the causation of observed normally distributed data, the "multiplex-elements" hypothesis. (See Ch. VI for further exploration of "elements.") A striking example of this multiplex-elements hypothesis in the sociological literature is Pemberton's discovery of the normal distribution of the rate of cultural diffusion of certain characteristics, such as the adoption of postage stamps by the nations, the membership in the national congress of parents and teachers, and the adoption of compulsory education by the States. (See S. 28, Ch. II, S. 5, Ch. X; S. 5, 29, Ch. XI.) This last provided more critical evidence on the multiplex-elements hypothesis. It was found that all the States together showed a non-normal distribution of dates of adopting compulsory public education. But the Northern States alone showed a normal distribution, as did the Southern States also when distributed by

\* For Eqs. 24a-e, Ch. V, see notes at end of the chapter.

themselves. The Southern States lagged about thirty years behind due to the negro problem. Within each section of the United States the culture was "homogeneous" in this respect, but between the two sections, the illiterate, ex-slave negro population constituted a factor dominating in influence over the multiplicity of factors tending to diffuse education more widely. In the case of other indicators (see S. 5, Ch. X, and S. 5, Ch. XI) of the diffusion of educational interests, the influences of the World War of 1914-18 and the economic depression of 1930-33 disturbed the normal ogive curve, but these disturbances were compensated for by the multiplicity of factors at work as soon as those major influences were removed. Further charting and fitting of curves to data of cultural diffusion should lead to identifying such normal trends while still uncompleted, and enable prediction as to their rate of completion (perhaps to S. 3, Ch. XI, for example).

## 2. THE EVOLUTION-TO-NORMAL HYPOTHESIS

Again, the type of distribution tells its stability. Distributions whose curves have infinite higher moments are unstable. The most stable type of all is the normal curve, all of whose moments are finite. This fact led Kelley to propose the hypothesis that evolution is a trend towards normal distributions. Arguing that evolution is a trend towards the stable, that the unstable being unstable does not endure, he suggests that the normality of a distribution should serve as a test of whether the distributing characteristic in a given population is evolutionally mature, or is still evolving and "on the make."<sup>14</sup> (See Ref. 35, pp. 147-150.) Kelley suggests that his hypothesis can be tested by comparing distributions of shells and skeletons of extinct and primitive species with surviving species.

## II. DISTRIBUTIONS OF INDICANTS OF DESIDERATA, $I_v:P$

### A. Definitions

Evaluated indicants and evaluated plurels have been noted in the two preceding chapters. They may now be discussed a little more adequately in describing their combination in valuated distributions, which are distributions of some desideratum.

A desideratum is anything, tangible or intangible, desired by people. It is any physical or psychic object satisfying a human want. A positive desideratum is defined as that towards which people behave so as to preserve or increase it. A negative desideratum is defined by behavior tending to decrease it. Desiderata are thus entirely relative to specified parties at specified times. They need not remain purely subjective matters confined to the fields of literature, ethics, and philosophy. They can be scientifically observed by any indicators of what people want.<sup>15</sup> Common types of indicators of desiderata may be money expenditure, time expenditure, verbal statements under conditions which reduce the probability of insincerity or rationalization, votes, attitude and opinion tests, behavior, such as attendance at meetings, writing letters, striking, etc., choices made in terms of experiences rejected in favor of alternative experiences, etc. The list may be greatly expanded. With research, technics for observing indicators of desiderata may be expected to become more comprehensive, reliable, and valid.

Whenever the indicators of what human beings desire (including negative desires and aversions) can be quantitatively expressed, a frequency distribution can be made. Such a distribution of a desideratum may be symbolized by the letter V, used either as the index or script, and serving as a subindicator.

$I_v^0$  (or  $V_v^0$ ) = value attributes, i.e., desired attributes,  $|_v$  in number (Eq. 25a, Ch. V)

$I_v$  (or  $V_v$ ) = value indicants, i.e., indicants of desiderata,  $|_v$  in number (Eq. 25b, Ch. V)

$\downarrow I : P$  (or  $\downarrow V : P$ ) = a distribution of persons according to a desideratum (Eq. 25c, Ch. V)

In dealing with value distributions and the plurels at their extremes the following terms may be used: "normals," are defined in this volume as parties falling between plus one sigma and minus one sigma, or the 68% which fall, in the normal distribution, between the 16th and 85th percentiles; "subnormals," those who are one sigma or more below the mean, i.e., the 16% from the 1st through the 16th percentiles inclusive; "super-normals," those one sigma or more above the mean, i.e., from the

85th percentile up; and "non-normals," the subnormals and supernormals combined. For more precise purposes, the plurel which is two sigma or more below the mean may be called "mininormals." In the normal distribution they will be the 2% in the first and second percentiles. Similarly the "maxinormals" are those who are two sigma or more above the mean, or the 2% in the 99th and 100th percentiles. The "abnormals," which combine the mininormals and the maxinormals, may often be more readily defined by percentiles in situations where indicants and standard deviations have not been developed. Thus the percent divorced, the percent insane, a mortality rate, etc., are defined by the points of a divorce decree, a commitment to an institution, death, etc., on conceivable continuums (for which indicants have not yet been adequately developed) of marriage adjustment, sanity, and health respectively.

These names for the class-intervals whose boundary points are multiples of sigma are general ones. In particular studies where boundary points that are objectively determinable mark off percentages other than the 2%, the persons in the extreme class-intervals may be conveniently referred to as "minimals," or "maximals," and defined by some non-integral value of sigma or by a percentage at the extreme which is not necessarily 2%.

For non-value distributions a similar set of terms for deviates having no implication of social desirability or undesirability may prove convenient, as defined below:

For a value distribution  $\downarrow I : P$  For any distribution  $\downarrow I : P$

Mininormals =  $\ast' \cdot 2\sigma, I : P$  = Minimates (Eq. 26a, Ch. V)

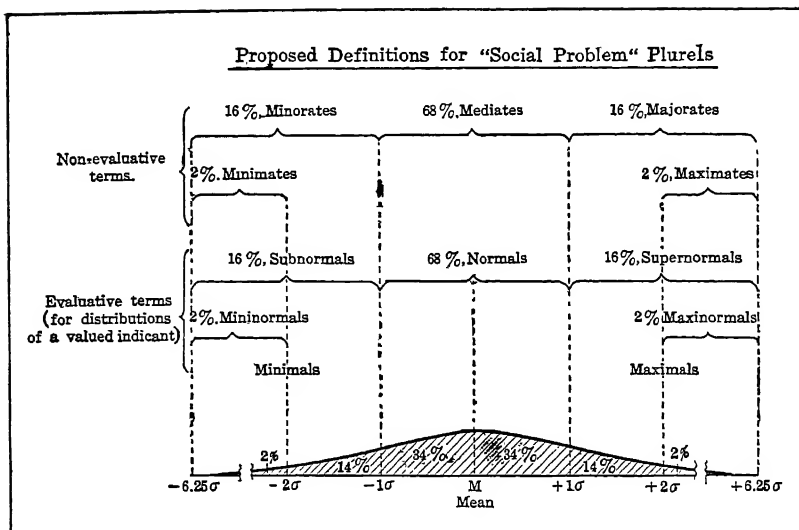
Subnormals =  $\ast' \cdot \sigma, I : P$  = Minorates (Eq. 26b, Ch. V)

Normals =  $\pm \sigma, I : P$  = Mediates (Eq. 26c, Ch. V)

Supernormals =  $+\sigma \cdot \ast', I : P$  = Majorates (Eq. 26d, Ch. V)

Maxinormals =  $2\sigma \cdot \ast', I : P$  = Maximates (Eq. 26e, Ch. V)

In these definitions the symbols, as usual in S-notation, state the limiting points in the point script (where  $\ast'$  always denotes the first or lowest limit, and  $\ast'$  the last or highest limit) of a particular class-interval,  $\downarrow I$ , for any indicant. A graph defining these terms follows herewith.



### *B. "Social Problems" or "Social Pathology"*

With these definitions, the whole field of social pathology can be more systematically treated. A "social problem," as currently used in the field of Sociology and social work, is definable as minimality, i.e., a defined extreme degree of deficit of some characteristic which is desirable to the population in which the social problem exists. The problem of poverty is an extreme deficit of the characteristic wealth, which, if valued by a population, thereby defines the poor as a social problem. In a monastery where wealth is not valued those lacking it in that plurel cease to be a social problem. The problem of crime is an extreme degree of negative law-abidingness, defined by the laws which formally state those of the values which are compulsorily enforced by the plurel of that legal jurisdiction. With another country or State with different values and laws crystallizing them, the criminal will be defined differently. The problem of the defectives is made up of those persons who lack in extreme degree the socially desired characteristics of health, sanity, and intelligence. The illiterates constitute a social problem plurel only in a population that desires literacy; divorcees are a problem wherever marriage adjustment is a positive desideratum; atheists may be a social problem in a Puritan community; and theists may be that problem in a Bolshevik community.

The systematic definition of social pathology is in terms of minimality. To define a social problem requires developing indicators of what a population desires. Extreme negative deviates from the mean of this desired characteristic are the social problem plurel. The chief types of human desiderata may be roughly classified as above, when classifying indicators (Ch. III) and plurels (Ch. IV). The indicators of institutionalized composite control environments are indicators of desiderata which people wish to control and may be referred to as valued indicators, I.. The consequent segmental plurels they form are plurels of specialists and their clienteles who desire the segmental desiderata.

### TYPES OF "MINIMALITY" <sup>16</sup>

<i>Extreme Negative Deviates on Some Valued Index</i>			
<i>Bernard's Derivative Institutional Control Environments (Modified) I.</i>	<i>Examples of Valued Indicators Distributing Characteristics in These Environments</i>		<i>Which Define Corresponding Social Problem Plurels, i.e., Minimals <math>a', -2a', (I) : P</math></i>
	$\mathbb{P}\mathbb{P} : (I)_v$	$\mathbb{P}\mathbb{P} : (I)_v$	
	<i>For Individuals in a Plurel, Chiefly</i>	<i>For Plurels in a Population, Chiefly</i>	
(1) Domestic	Marriage adjustment scores (Burgess and Cottrell)	Divorce, marriage, fertility, illegitimacy rates	Divorcees, prostitutes, illegitimates
(2) Medical	Temperature, blood pressure, and other physiological tests, and psychoneurotic tests	Morbidity and mortality rates. Institutional commitments. Public Health Appraisal Forms	The sick, the defective, blind, crippled, insane, etc.
(3) Economic	Income. Days per year employed	Average income, percent unemployed	The poor, the unemployed
(4) Political	Attitude tests. Party membership. National citizenship plus indicators of degree of war participation	Crime rates. Political arrests and prisoner rates. Rates of war damage in bereaving, maiming, impoverishing, exiling	The criminals. Political refugees. War-caused dependents
(5) Racial	Skin color, hair type, cephalic indices	Anthropometric indices	Inferior races—Bushmen, Pygmies, etc.

TYPES OF "MINIMALITY" <sup>15</sup> — (Continued)

<i>Extreme Negative Deviates on Some Valued Index</i>			
<i>Bernard's Derivative Institutional Control Environments (Modified)</i> <sub>17</sub>	<i>Examples of Valued Indicators Distributing Characteristics in These Environments</i>		<i>Which Define Corresponding Social Problem Plurels, i.e., Minimals <math>a', -2\sigma, (I)_v : P</math></i>
	$\underline{P}_1 : (I)_v$	$\underline{P}_2 : (I)_v$	
	<i>For Individuals in a Plurel, Chiefly</i>	<i>For Plurels in a Population, Chiefly</i>	
(6) Educa- tional	School achievement, Tests of informa- tion or skill	Illiteracy rates, An- nual books per capita. Educa- tional budgets	The illiterate. The ignorant or "backward" peoples
(7) Recrea- tional	Variety of leisure activities. Rating in a particular leisure activity	Per capita public ex- penditure for re- creation. Weighted indices of partici- pants $\times$ occasions $\times$ types of recre- ation	The overworked and underprive- leged classes (not necessarily the lowest incomes)
(8) Esthetic	Artistic appreciation tests (e.g., Sea- shore's Musical Talent). Contests for artistic pro- duction	Indices of institu- tions for art, music, gardening, etc.	The offensively vul- gar, vandals
(9) Ethico- Reli- gious	Attitude indicators, Participation indi- cators. Character tests (May)	Church membership and attendance rates. Institu- tional indices. (Chapin)	The atheists or anti-religious. The vicious
(10) Linguis- tic	Standardized lan- guage achieve- ment tests	Average scores on language tests. Percent speaking a language	Foreign language minorities
(11) Com- munal	Percent of communi- ty's culture traits participated in. So- cial distance scores	Social distance scores between communities	The hermit or iso- lated. The ostrac- ized

## C. "Social Therapy," or Ways of Decreasing Subnormality

Since social problems are negative values, people consciously strive to reduce them, at the same time as they often unconsciously through ignorance, conflict of interests, etc., behave in ways which perpetuate the problem. There are three general

methods of reducing subnormality. These three methods are defined by the first three moments of distributions.

### 1. PROGRESS, ${}^M(I)_v$

The first method is to raise the whole population, to increase the mean of its valued indicant or index. This method moves the whole distribution towards larger amounts of the desired characteristic. (For a controlled experiment graphing the distributions of such progress in hygiene in experimental, control, and demonstrational villages, see S. 30, Ch. II.) It increases the mean indicant,  ${}^MI$  (and all person-indicants,  ${}^PI$ ) by the amount progressed without otherwise changing the distribution. Everyone has a constant increment with no change in their relative standing. Its symbol is:

${}^M(I)_v$  = progress to date, the mean of an index valued by a population (Eq. 27, Ch. V)

The dynamic aspect of this method, the change in the mean with time, is the process of "progressing" and will be taken up in Chapter X, when dealing with the time sector. This method is defined by an increase of the first moment. (See Eq. 8, Ch. V.)

### 2. EQUALIZATION ${}^E(I)_v$

The second method is that of equalizing by drawing both extremes of the distribution in towards the mean. To reduce extremes of wealth, redistributing it more equally would be an economic form of such equalizing. To abolish the privileges of the nobility and enfranchise the humblest citizens would be a political form of equalizing. Democracy in various forms aims to equalize. This method is a shrinkage in dispersion and is measured by a decrease of the standard deviation. The stage or degree of equalization existing at any date is symbolized by:

${}^E(I)_v$  = equalization, the dispersion of an index valued by a population (Eq. 28, Ch. V)

(Again the dynamic aspect, the process of "equalizing," is discussed more fully in Chapter X. For an example of the opposite, disequalizing, in incomes, see S. 26, Ch. XII.)

This method, then, is defined by a decrease of the second moment (see Eq. 9, Ch. V).

3. AMELIORATION,  $sk(I)_v$ 

The third method of reducing subnormality is to decrease the minimals' class. To enrich the poor, heal the sick, prevent divorces, teach the illiterate, convert the atheists, re-employ the unemployed, reclaim the drug addict, and to prevent or to uplift the minimals of any kind, would be amelioration (in any population where the plurels listed in this sentence were negatively valued). This method draws in the lower extreme of the distribution towards the mean. It shifts the skewing of the curve from one side towards the other. It is measurable by a positive increment in the coefficient of skewness. Its static aspect, the degree of amelioration achieved up to a given date, is symbolized by:

$Sk_v$  = amelioration, the amount and sign of the skewness of the distribution of a population on a valued index (see Eq. 10, Ch. V). (Eq. 29, Ch. V)

This method, then, is defined by a positive increase of the third moment.

For an application of these concepts it may be suggested here that the ideal of capitalism are methods (1) and (3), i.e., to reduce economic pathology by general progress and amelioration (as defined above), while the ideal of communism are methods (1) and (2), i.e., to solve a population's economic problems by general progress and equalization (as defined by the equations above).

*D. An Hypothesis of a "Natural Range"*

A sociological issue is presented by the dogma of capitalism, that individual differences in initiative and ability justify the existing degree of dispersion of wealth, and the opposite dogma of communism, that complete justice lies in equality of economic income (whatever inequalities there may be in psychic income). The issue appears in its political form in the current opposition of dictatorial systems to the more equalitarian democratic systems.

A straddling middle ground, such as in a philosophy of economic inequality coupled with political equality, seems to be increasingly difficult to maintain. The issue is becoming increasingly emo-

tionalized. Can sociologists contribute any light tending to resolve the conflict? What theories are there which might tend to solve the problem by the rational technics of science rather than by force?

Upon this issue, an hypothesis is offered below, which, while of no practical significance immediately, may ultimately, if verified, and applied, contribute to solving the problem.

The issue can be stated, theoretically, as the degree of dispersion desired (including assented to) by a population. For given characteristics (chief of which are income and standard of living, political rights and power) what amount of dispersion shall be tolerated or striven for? Shall it be slightly more or slightly less than the current degree, or radically more or radically less approaching the maximum or the minimum? Shall the policy be conservative or liberal, fascist or communist? Of course the issue thus stated is oversimplified. Actually it is complexly overlaid with other issues more or less correlated. But it is submitted that at least a major element in the complex of issues that is increasingly drawing nations into opposing camps is the question of the optimal dispersion within a nation (and also between the nations), which, of course, may be different for different characteristics.

Furthermore, each nation's attitude to the current distribution of raw materials, markets, colonies, population outlets, military power, prestige, etc., is its value judgment for that characteristic, and determines what it considers to be a more nearly "optimal" dispersion. But interests conflict, and the definitions of the "optimal" for any one characteristic may be as many as the parties whose desires define it. Is there any criterion of the "optimal" dispersion which, by its scientific demonstrability, might gradually become accepted and be collectively acted upon? Our hypothesis on this point is that a tendency to a "natural range," to a "normal" amount of dispersion, exists and can progressively be discovered<sup>17</sup> by research.

Suppose, for example, a comprehensive research were made to collect all the distribution curves on record for any plurels on all characteristics, biometric, econometric, anthropometric, sociometric. What percentage are normal curves (i.e., have  $\beta_1 = 0$ ,  $\beta_2 = 3$  within standard error limits)? What is the maximum range from lowest to highest limit in each distribution as meas-

ured by sigma units? The normal probability table tells us that a range of six sigma will be exceeded about three times in a thousand instances, and that a range of  $12.5\sigma$  will not be exceeded once in two billion times. (Ref. 37.) As the population of the world is estimated to be not far from two billion, this means that in a normally distributed characteristic not one person will be found in the whole world beyond the range of twelve and a half sigmas, i.e., beyond  $6.25\sigma$  from the mean.<sup>18</sup>

$\pm 6.25\sigma$  I : P = .8 = probable number of living persons beyond a range of  $12.5\sigma$  in any normally distributed characteristic (Eq. 30, Ch. V)

Let us, then, adopt this range of  $12.5\sigma$  as the upper limit of the possible range for any normally-distributed human characteristic. The sample here is maximal, as it is all living human beings. Now, do observed distributions ever show ranges greater than  $12.5\sigma$ :

- a. when the distribution is well fitted by a normal curve?
- b. when the distribution is not normal?

The ranges found should themselves be tabulated in a frequency distribution and the mean range, the maximal one, and the percentage exceeding any defined range, such as  $12.5\sigma$ , should be determined. Those exceeding  $12.5\sigma$  should be carefully studied as to size of sample, artificial or unnatural factors in the situation, etc. Our hypothesis of a "natural range" is that  $12.5\sigma$  tends to be the upper limit of range of human characteristics that are naturally, i.e., non-culturally, determined. This hypothesis can be crucially tested (though the labor involved is great) and its degree of truth, as shown by the indices of the distribution of ranges of distributions, can be ascertained. The hypothesis implies that culturally determined distributions may be expected sometimes to show ranges exceeding  $12.5\sigma$ , but that characteristics which are least affected by human culture, which are most natural or hereditary, will not exceed that limit. (The degree of cultural determination will have to be assessed and a rating, or other scale, for it may prove extremely difficult to construct.) Thus, distributions of IQ's from reliable and well-calibrated intelligence tests seldom exceed a range of  $8\sigma$  sigma. (See for ex-

ample S. 21, Ch. XII.) But economic incomes in our capitalistic culture show no such limit. Incomes of a million dollars a year (which are well below the maximum incomes) constitute a range from the smallest incomes, of some 2000  $\sigma$ . (See S. 20, Ch. XII. In the graph the dispersion has been enormously shrunk by using logarithms of dollars as the distributing unit.) No calibrated tests of ability have ever, to the author's knowledge, shown such a tremendous dispersion as this. Do not such findings, if further corroborated, throw doubt on the claim that the current largest salaries and incomes are justified because they are proportioned to individual differences in "initiative" and "ability"? On the other hand, the scientific evidence reported by psychologists disproves the doctrine that all men are created equal, as it is probable that every human trait, if measured finely enough, would show a dispersion. May not a solution between these opposite extremes of doctrine lie in finding some reasonable limits of range, and, by social convention, shaping our cultural characteristics proportionately? The problem of how to enforce such conventions by law or collective action can be safely postponed for many years to come, as the prerequisite scientific job will require decades of work by thousands of able, well-trained, and well-supported researchers. There are great technical difficulties in devising indicants for intangible characteristics, in applying them adequately to whole populations, in analyzing the resultant distribution curves, and in gradually establishing a consensus among scientists as to the emergent findings which can be a factual base on which public opinion and mores, and eventually laws and institutions can be built.<sup>19</sup>

*E. A Tension Theory of Societal Action, 'E'*

Thus far the objective aspect of values and the objects valued, have been dealt with. But the subjective aspect of the evaluation, the intensity of the desire in people, which determines what is a desideratum, is the more important sociologically. For it is the desires, the wants, the motives of people in certain situations that create human desiderata and determine individual and societal behavior. In order to deal more scientifically with both aspects (the object valued and the desire of the evaluating person) a theory, or system of hypotheses, is hereby offered.

## 1. DEFINITIONS AND ASSUMPTIONS

Suppose that behavior is summarized in the statement: "People desire desiderata," meaning that people continually want objects of psychic and physiological satisfaction. This somewhat circular statement that people desire what they desire becomes useful only if we have an equation in which all terms but one are determinable from observation, enabling the finding of the solution for the unknown quantity. To accomplish this, let  $P$  be the number of people concerned with a particular desideratum; let  $D (= {}^M I_p)$  be the average intensity of their desire for that desideratum; and let  $V (= I_v)$  be the quantity of units of the desideratum that are available for supplying that desire.

$V : P : D =$  a value with its corresponding evaluating persons and their average intensity of desire (Eq. 31, Ch. V)

Thus, twenty athletic competitors ( $P = 20$ ) want one of three prizes ( $V = 3$ ) with an intensity of desire measurable in standard deviation units (say  $D = .3\sigma$ ) on some scale such as (for a behavioristic index), the number of hours each will spend in training for that contest, or (for an attitudinal index) an attitude test of endorsing such statements as, "I would rather win this prize than graduate," etc.

In explanation of the term, "*desire*" should be interpreted broadly as all sorts of "wishing," "hungering," "interest," "striving," or other concepts of conscious or unconscious inner motivation. Desire may be evidenced in its covert state as attitudes or in its overt state as behavior, and this may be highly mechanistic (as in thirst) to highly telic (as in an ideal of a utopia, where the idea is physiologically a current or tension in nervous pathways and, therefore, just as much a prior cause of behavior as the thirst is to drinking). "Desire," as all inner motivation to behavior, becomes of absolute zero intensity only in death, though it may temporarily be in suspense, as in sleep. Broadly, it is the total inner states (including experience) of the organism determining the response upon stimulation, as indicated in the modified stimulus-response formula:

$$B = (0 ; S) \quad (\text{Eq. 32, Ch. V})$$

behavior is some function of the organism and the stimulus-situation.<sup>20</sup>

The desideratum is the philosophical "value" more than the economic one. It means the object, or objects, of desire which may be anything from mystic communion with the Absolute to bread and butter. Although often highly abstract and symbolic, it may be external to the individual, and therefore, more objectively observable than the desire which is entirely subjective and can only be inferred from speech and other behavioristic indicators. Thus, even where the desideratum is "prestige," it is observable in the glances, acts, words, and tones of associates, etc. In economic realms it is the goods or services that satisfy wants. But in other realms it may be a political office, a mate, graduation, fame, or anything one sets one's heart upon and behaves in such a way as to get more of it or keep from losing it. Often it will be a whole, a unitary and indivisible qualitative entity, such as "winning" a contest, and then  $V^0 = 1$ . Deciding on the units, or entities, whose number is the quantity  $V$ , is the central problem of measurement in this theory. In many situations it is determinable, and in many others should yield to competent research. The theory is fully applicable only when  $V$  is determinable. Such situations are rare in the present stage of methodological development in Sociology.

Now the fundamental fact is that human desires usually exceed the values available for satisfying them,  $PD$  is greater than  $V$ .

$PD > V$  = human desires tend to exceed the available desideratum  
(Eq. 33, Ch. V)

Man wants more of what he wants, whether it is a positive getting something or a negative avoiding something, or whether it is the mere further living in order to continue experiencing. When *all* desire goes, death ensues; here  $D = 0$ . Hence  $PD$  vanishes mathematically in the equation and physically in death. The inequality (Eq. 33, Ch. V) describes in cold mathematical symbols the urge to activity in all living creatures. In special disciplines this inequality is described in terms connoting associated facts in that field, such as "struggle" for existence in biology, a "strain toward equilibrium of forces" in bio-chemistry, a "drive" in psychology, a "group interest" causing a societal process in

sociology, etc. The mathematical statement is here preferred as the most definite and universal, since each discipline may evolve for each term of Eq. 33, Ch. V, its most appropriate units, the interrelationship of which, expressed by Eq. 33, Ch. V, may be valid to all.

## 2. THE TENSION OR EQUILIBRIUM EQUATION

Next, to convert the relation Eq. 33, Ch. V, into an equation, introduce a coefficient E defined by:

$$PD/V = E, \text{ the equilibration (or "tension") index} \\ (\text{Eq. 34a, Ch. V})$$

or in equivalent, and more symmetric form:

$$PD = VE = \text{the simple tension theory of societal action} \\ (\text{Eq. 34b, Ch. V})$$

The ratio E measures the degree to which the total desire of a population exceeds the available amount of the desideratum in question. Since it thus measures the degree of equilibrium between a desire and its satisfying desideratum, it may be denoted by E, the equilibration ratio. Various other possible ways of thinking of it are, as a "coefficient of value," or "worth" of each unit of V, or again as a desire-satisfaction ratio in societal fields, or as an index of social "tension." It is a ratio of demand to supply in economics (before correcting the statement of that relation by the concept of price). A simple mnemonic statement has been suggested as: "Peoples' desires and the objects valued are to be equalized."

Before clarifying these concepts with numerical illustrations from various social disciplines, it should be explicitly noted that D is the average desire in a population whose intensity of desire may vary from person to person in some sort of a distribution.

$$D = \Sigma(I)_D/P = \frac{\text{sum of indices of desire}}{\text{number of people}} \quad (\text{Eq. 35, Ch. V})$$

When this average is multiplied by the number of persons, as in Eq. 34a, Ch. V, it leaves:

$$E = \frac{\Sigma I_D}{V} = \frac{\text{total desire}}{\text{total satisfaction}} \quad (\text{Eq. 36, Ch. V})^{21*}$$

\* For Eqs. 37 and 34c, Ch. V, see notes at end of chapter.

It should also be noted that for the case where the population is one person,  $P = 1$  and  $E = D/V$ , a person's tension is the ratio of his desire for a value to the quantity of the value available.

### 3. NUMERICAL EXAMPLES

#### *a. From political science*

For a political case of the theory, consider nationalism, the intense desire for the desideratum "national aggrandizement." Such a unitary qualitative desideratum is denoted by an exponent of zero.  $V^0 = 1$ , and so  $E = PD$ . The tension of the nation towards this desideratum varies directly with the number of people in the nation and with the average intensity of their nationalistic desires.

Democracies have been applying this equation (Eq. 34, Ch. V) for centuries. Except where a person has several votes (weighted by property held, etc.), or in proportional representation, every person is classed as having 0 or 1 units of intensity of electoral desire. The criminals, non-citizens, and those not interested enough to register get a 0; the rest all get 1, simply because individual differences could not be measured reliably, so the effort was given up (when aristocracy or the hereditary system of weighting individuals was abandoned) and everyone was considered equal, as a rough but closer approximation to ideal weighting. In a situation, then, where 1,000,000 voters are electing 20 members of Parliament, the coefficient of value of each seat is 50,000 (in terms of votes).

$$\begin{array}{ll} P = 1,000,000 & V = 20 \\ D = 1 & E = 50,000 \\ 1,000,000 \times 1 = 50,000 \times 20 \end{array}$$

(Eq. 38, Ch. V)

This measures precisely the common opinion that the worth of an office varies according to the size of its electorate, the mayorship of New York City outweighs that of a small town; the presidency is more coveted than a governorship. This is the worth of the office. For the relative worth to the electorate of each candidate the equation is applied in turn to each candidate as the unitary value. The number of votes cast for him measures

his official worth in a democracy relative to the other candidates. The sum of the worths,  $E$ , of the candidates, equals the worth of the office. This varies from one election date to another, as the varying turn-out of voters indicates.

Thus, in a democracy, elections are a crude device to measure and apply Eq. 34, Ch. V.

*b. From education*

For an application of the formula in the field of education, consider the value of a college degree to the population identified as "entering Freshmen." If about 50% of them will graduate, in the experience of that particular college, and there are 500 Freshmen, then  $V$ , the number of coveted degrees achievable, is about 250. As one rough index of desire,  $D$ , take the number of years of effort of the sort defined by passing courses and fulfilling normal expectations of campus life.  $D$  is a little less than 4 years—depending on how early that college eliminates the weaker students—say  $D = 3$  in units of years of "collegiate effort."<sup>22</sup> Then in these years of "crude effort" units

$$\begin{aligned} P \times D &= V \times E \\ 500 \times 3 &= 250 \times E \\ E &= 6 \end{aligned} \quad (\text{Eq. 39, Ch. V})$$

This "worth" of 6 units for the degree in that college<sup>23</sup> may be compared with the situation in another college where it may be "worth" more, since, for example,

$$P = 500, D = 1.8, V = 300 \text{ and, therefore, } E = 3 \quad (\text{Eq. 40, Ch. V})$$

In the second case the percent failing to graduate is smaller and they are eliminated earlier in their course, resulting in smaller "societal tension,"  $E$ . The "worth" or "tension" towards a degree for a group is thus seen to vary directly with the number of years of effort required, and inversely with the probability of passing.<sup>24</sup>

*c. From biology*

For a biological example take the Malthusian principle. At any time people,  $P$ , have an average intensity of desire,  $D$ , for

food and wherewithal to live,  $V$ . The ratio of the total desire to the available maintenance, in whatever units these are expressed, is  $E$ ,<sup>25</sup> the coefficient of tension or biological price of one share of maintenance.

Now as time goes on,  $P$  tends to increase geometrically by births and  $V$  tends to lag behind in arithmetic increase (according to an early version of the Malthusian principle); with the resultant increase of  $E$ , the biological price of each share of maintenance.

$D$ , here, is not as narrow a concept as conscious desire to live, but is the broad total desire of the organism, its total urge to adapt. Since individuals vary in this ability to adapt,  $D$ , here, is as always the average of a distribution,  $\frac{\sum I_D}{P}$ . As  $E$ , the biological price (in terms of difficulty of effort and amount of intelligence required for each share of maintenance) increases, it passes the maximum capacity of the inferior individuals who are eliminated in the struggle to survive. They have not what is needed to get a unit of  $V$ , therefore they get no  $V$ ; they die; they drop out of the equation, thereby reducing the left-hand member  $PD$  and tending to bring it into balance with the right-hand number,  $VE$ . This selection is more drastic and swifter when  $E$  is large; the struggle implied in the inequality (Eq. 33, Ch. V) is greater then; but the selection is less rigorous (i.e., a smaller percent of  $P$  is eliminated) when  $E$  is small.

Biologically the problem of balancing Eq. 34, Ch. V, i.e., keeping  $E$  small, can be met in only three ways:

1. Reduce the rate of increase of  $P$ —naturally by the death of the less fit, or artificially by birth control, migration, etc.
2. Decrease  $D$ , lower the standards of living—as in the river valleys of India and China.
3. Increase  $V$ , the supply of food and things needed for human life—by scientific technology, and collective organization.

Now consider the case of a society (as in the Occident in the century since Malthus) when  $V$ , the food supply, by reason of scientific technology, grows faster than the population. (See S. 32, Ch. II.) If our hypothesis (Eq. 34, Ch. V) is true, one (or more) of three things will then happen to compensate for the increased food supply:

1. The population will tend to grow faster with less elimination of the unfit; or
2. desires, the psychological standards of living, will rise; or
3. the social tension, the urge to satisfy unsatisfied desire will decrease, i.e., society will become more leisurely and less strenuous.

Thus, whatever the facts as to the Malthusian ratios in particular societies, Eq. 34, Ch. V serves to express the interrelation of population, production, and motivation factors, and to help predict the consequences of changes in some of the factors upon the other factors.

*d. From economics*

In Economics, theories of value have been developed with great refinement. Essentially, economic value rests on scarcity (otherwise the object has mere "utility"), which is another way of saying that Eq. 33, Ch. V is an inequality, that potential demand exceeds supply. To provide a common unit of economic values (i.e., goods and services) value is related to money, giving rise to the ratio "price." Without multiplying examples wearisomely, it may be briefly pointed out that the economic value (our E) varies directly with the number of competing buyers (our P) and with their desire in money offered (D), and inversely with the supply (V) of the goods. This is as stated in Eq. 34a, Ch. V without the refinements required in Economics by the volume and velocity of circulation of money and credit which also affect price, and without stating the further complications of monopolies and other restraints on free competition.

$$E \text{ (economic value)} = \frac{PD \text{ (total money offered, demand)}}{V \text{ (supply)}} \quad (\text{Eq. 41, Ch. V})$$

*e. From religion*

For a contrastingly intangible field, consider in religion the value to mystics that may be stated as "closer communion with God through prayer." Theoretically there is an infinite supply of V, but actually, as drawn upon by human beings there cannot

be more per capita shares of such experience (whatever their size) than there are human beings, so that  $V$ , the number of average-sized shares, as realized, cannot exceed  $P$ . We may take the average share as the unit  $\frac{V}{P} = 1$ . Then  $V = P$ . Then  $E$ , the worth of such communion to the population as a whole, is equal to the average intensity of desire for it by the individuals in that population.

$D = E$ , i.e., tension equals average desire, when

$P = V$ , as in unlimited desiderata (Eq. 42, Ch. V)

f. *From philosophy*

One philosophical implication of this theory is, that all interaction (and all behavior) has for its objective the minimizing of the tension  $E$ , in Eq. 34, Ch. V. The trend of behavior is to make  $E$  approach 1 as a limit.

$E \doteq 1$  = the goal of behavior (Eq. 43a, Ch. V)

or  $PD \doteq V$  (Eq. 43b, Ch. V)

When  $PD = V$ , the available values equal the total desire, a perfect equilibrium is reached. But since  $D$  grows within us as life and mental life itself, and since  $P$  grows with the procreative sex urge, there is forever an increasing of  $PD$  and a resulting strain, an inequilibrium which drives us to action making us strive forever to win more desiderata, to come nearer the elusive equilibrium.

Three alternative philosophies of life have grown up to ease this strain, to bring Eq. 33, Ch. V into equilibrium, to realize Eq. 43, Ch. V. In the Orient, at its extreme in Buddhism, renunciation of desire is the solution. Whittle away desire "for it is illusion." Decrease  $D$  in Eq. 34, Ch. V in this life and in future incarnations, until at last the elect attain Nirvana, the blessed state of absence of all desires. In the Occident, at its extreme perhaps in America, work to satisfy desire is the solution. If you want something, "go get it," or at least try! Increase  $V$  in Eq. 34, Ch. V with invention and mass production so that more and ever more of our desires can be gratified.

Thus, the objective and problem of living, stated mathemat-

ically <sup>26</sup> in Eq. 36, Ch. V is solved typically in the East by more accommodation.

$$D \doteq 0 \quad (\text{Eq. 44, Ch. V})$$

and in the West more by organized effort in co-operation.

$$V \doteq PD \quad (\text{Eq. 45, Ch. V})$$

This, of course, is an extreme tendency in each region with enormous overlap of the two groups.

A third philosophy offering a way to reduce the tension, E, is that of Nietzsche, who advocated the ruthless elimination of the less fit, i.e., to make P decrease. This seems the philosophy underlying Fascism. Eq. 34, Ch. V shows that these three are the only ways to reduce the perennial tension—*increase the production of desiderata, decrease desires, or, for desiderata that are scarce, decrease the number of sharers.*

#### 4. THE SPECIAL CASES OF NEGATIVE, QUALITATIVE, AND UNLIMITED DESIDERATA

In this tension theory, three special cases should be studied. One is the case of negative desiderata or aversions. In this case the desideratum, V, although referred to as negative is treated with its normal sign, i.e., positive if there is any amount of it greater than zero, and negative when less than zero, as in debt. The descript may have a minus sign attached to denote that it is negatively valued. The desire, D, becomes negative in aversions passing through zero when the attitudes are neutral, neither desiring nor averting. Consequently, as the population, P, is positive, the tension, E, takes on positive or negative values as a tension *towards* more of a positive desideratum and a tension *away from* a negative desiderata.

$$P(-D) = V_{-V}(-E), \text{ the case of a negative desideratum} \\ (\text{Eq. 46, Ch. V})$$

Another special case is when the available desideratum is unlimited or at least exceeds indefinitely the desires of the population in that desideratum field (i.e., desiring that desideratum). Economic value theory has dealt with scarcity of desiderata.

Desiderata that are abundant must also be included in value theories. This is a problem for research. At present it is handled as a working hypothesis, as in the illustration of an unlimited desideratum in religion. Although the available quantity of the desideratum is potentially unlimited (such as a "belief") yet actually, there is only a finite amount used or experienced. If this finite amount is undeterminable, call it  $V$ , and deal with it as an algebraic unknown.

Then dividing  $V$  by  $P$  gives the average, or per capita, share utilized—in whatever imaginable but undeterminable units such "shares" may be expressed. The conceptual average share becomes a unit which, by definition of the mean, can be substituted for any one person's share and will give the same result *in calculations concerning the plural*. The number of these average shares is the number of parties sharing, so that :

$$\begin{array}{l} V = P \text{ (or } \mathbb{P} \text{ if plurals are the unit) (Eq. 47a, Ch. V)} \\ \text{and} \quad D = E, \text{ in the case of an unlimited desideratum} \\ \qquad \qquad \qquad \text{(Cf. Eq. 42) \qquad \qquad \qquad (Eq. 47b, Ch. V)} \end{array}$$

For when Eq. 47a is substituted in Eq. 34b, Ch. V the  $V$  and  $P$  cancel each other out, leaving Eq. 47b, which states that the average desire for the unlimited desideratum equals the societal tension towards it. Tension and average desire become the same thing when there is no limiting scarcity of the desideratum to alter the balance. But, for a constant desire, if the available desideratum became scarce, tension would go up; competition to get it would become keener; tension and average desire would no longer be equal.

Examples of this treatment of unlimited desiderata are numerous. The economists, stock example of free goods, such as the air we breathe, is one. Friendship is another potentially unlimited desideratum. The average desire for friends in a plural determines the tension towards friendships, the urge to develop friendships, the societal worth of friendships to that plural.

This treatment of an unlimited desideratum is closely related to treating it as a qualitative unitary desideratum,  $V^0$ . (See the example of national aggrandizement above.)

Thus, "friendship" may be taken as a collective unitary desideratum,  $V^0$ . Substituting it in Eq. 34b, Ch. V and then into

Eq. 36, Ch. V, and remembering that any quantity with a zero exponent is one, gives:

$$V^0E = E = PD = \Sigma I_D, \quad \text{the case of a unitary qualitative desideratum} \quad (\text{Eq. 48, Ch. V})$$

This states that the societal tension towards any qualitative unitary desideratum is the sum of the desires of the people concerned with that desideratum. Thus, the tension towards "friendships" (a plural noun) is the average desire<sup>27</sup> of the people for friendships, while the tension towards "friendship" (a collective singular noun) is the sum of those desires.<sup>28 \*</sup>

For the cases of unlimited, or of qualitative unitary desiderata little is gained by the concept of "tension," E. The average desire, D, is simpler and more direct. But E becomes useful for scarce or limited desiderata and for systems of desiderata, and the tension theory equation (Eq. 34b, Ch. V) is general to them all.

## 5. SYSTEMS OF DESIDERATA

To make this tension theory more adequate in describing complex societal situations, it should be noted that any of the three independent factors, P, D, or V, and the dependent factor, E, may be determined on different groups, at different dates, and with respect to differing desiderata.

The case of different dates or periods,  ${}_t|$ , involving the time sector will be reserved for Chapter X on Change. For the different desiderata in a situation the class script, as usual, specifies their number and kind,  $|_v$ . For the different parties desiring those desiderata, the descripts, as usual, specify their number and kind,  ${}_p|_p$ . Each party has its own intensity of desire, D, for each desideratum, of course. The aggregation of these is symbolized by the matrix (using cross-scripts) as:

$${}_t:{}_p(PD = VE)_{v:p}^s, \quad \text{the matrix equation of the tension theory for distributions of valued indicants at different times.} \quad (\text{Eq. 49, Ch. V})$$

The scripts, of course, must be changed appropriately by the usual rules of S-notation for cases where periods instead of dates are involved ( ${}_t|$  becoming  ${}_t|$ ), and where the plurals are defined

\* For Eq. 47c, Ch. V, see notes at end of chapter.

by attributes other than the desideratum,  $|_p$ , and where plurals are substituted for persons as the unit of population,  $P$  (i.e.,  $P$  shifts to  ${}_pP$ ). Wherever the scripts denote the singular case of one date, one plural, and one desideratum of primary factors, we have the singular case of the tension equation which is Eq. 34b, Ch. V.

To summarize any situation, an additive combination of the aggregated entities may be used as a first approximation. Weighting, or further refinements in combining, can be worked out as needed in particular problems where a system of desiderata among many plurals on different dates may be involved. First, note that the factors of one equation are defined by the desideratum, not by the plural. The plural determines the intensity of desire for the desideratum. There may be many plurals and as many  $D$ 's for one desideratum in one equation with one tension, but if more than one desideratum is involved the situation becomes a matrix of as many equations and as many tensions as there are desiderata.

The general formula for a sum of the aggregates in the tension theory is:

$$z_t: \sum_p^t (PD = VE)_{\sum_v: \sum_p}^{\sum_v: \sum_p} \text{ the summational tension theory (in S-script notation) (Eq. 50a, Ch. V)}$$

$$\text{or } \sum_{111}^{tpv} (PD)^e = \sum_{11}^{tv} (VE)^e \text{ (in conventional summational notation) (Eq. 50b, Ch. V)}$$

This may be written in expanded notation for one date ( $t = 1$ ), as an equation of two matrices of order  $p \times v$  and  $1 \times v$ , as follows:

$$\left\| \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right\| \begin{array}{c} (PD)' + ''(PD)' + ''''(PD)' + \dots + {}_{p'}(PD)' \\ (PD)'' + ''(PD)'' + ''''(PD)'' + \dots + {}_{p'}(PD)'' \\ (PD)_{v'} + ''(PD)_{v'} + ''''(PD)_{v'} + \dots + {}_{p'}(PD)_{v'} \end{array} \left\| = \left\| \begin{array}{c} VE' \\ VE'' \\ VE_{v'} \end{array} \right\|$$

(in expanded notation) (Eq. 50c, Ch. V)

If written without the plus signs, each row of the left-hand matrix might be a frequency distribution in which the intensities of the desire,  $D$ , are the distributing units at each of the class-intervals of which there is a frequency, a  $P$ , of the distributed units. As written with plusses, the row is summed, expressing the weighted total intensity of desire for the whole population. Of course, the  $P$  value in any cell may be zero, reducing the  $PD$

in that cell to a nul entry, and so the matrix may actually be far more "patchy" than the neat rectangular arrays portrayed above.

For an example of such a complex situation of a value system, consider the Russian Five-Year Plan. The Russian people, P, officially desired  $I_v$  diverse desiderata (see S. 31, Ch. II; S. 22, Ch. XI; S. 23, Ch. XII), such as  $V'$  = tons of coal mined annually,  $V''$  = bushels of wheat produced annually,  $V'''$  = horsepower of electricity available at the end of five years, etc. These may be thrown into some common unit, such as percentages of increase from some basal period. If the national intensity of desire were measured in some crude index, such as the number of days of labor to be devoted to the achievement of each desideratum, the equation becomes determinate with its terms known from observation. The fact that regional Soviets (plurels) had different intensities of desire and differing schedules of producing their portion of each total desideratum V is allowed for in the summation with respect to p in the columns for the  $|_p$  Soviets. Note that the columns in this illustration are qualitatively different plurels,  $|_p$ , and not class-intervals of a frequency distribution. Of course it is possible and simpler to write an equation for each desideratum for each Soviet. But since all the groups and all the desiderata were parts of a unified system, their combination in some such way as in the matrices (Eq. 50c, Ch. V) would seem a truer mathematical description of the situation.<sup>29</sup>

For another example of a system of tension equations consider the process of making a budget whether personal, institutional, or national. Here the various items of the expenditure half of the budget, are the v different desiderata, the quantity of each of which is measurable in dollars (or alternatively in percentages of a limited total of dollars). The aim of the budget making is to equalize the tension towards each of the desiderata. This may be represented in the chain equation:

$$E' = E'' = E''' = \dots = E_v, \quad (\text{Eq. 51a, Ch. V})$$

$$\text{or} \quad \frac{D'}{V'} = \frac{D''}{V''} = \frac{D'''}{V'''} = \dots = \frac{D_v}{V_v} \quad (\text{Eq. 51b, Ch. V})$$

(where P cancels out as it is a constant throughout).

As soon as the budget maker realizes that the intensity of desire (such as  $D'''$ ) for one desideratum is so large in comparison

with the dollars allocated to it ( $V_{///}$ ) as to make that ratio exceed the others, this excess is reduced by taking dollars from desiderata showing subaverage ratios and reallocating these dollars to  $V_{///}$  till  $D_{///}/V_{///}$  is equalized with the other ratios. The process of determining  $D$  is usually crude. In the case of one person, he may not verbalize  $D$  in any units but merely have a total feeling that  $V_{///}$  ought to get more dollars and that some other item ( $V_x$ ) can be pruned a bit. In the case of a group, speakers will voice a desire, a canvass may roughly gauge its extent, or a majority vote on an amendment will formally show it. These and other technics roughly tend to actions equalizing the tensions for the various budgeted items, i.e., equilibrating the system.

Further development of this matrix tension theory will be reserved for treatment with the time index included when discussing societal forces and control in Chapter XI.

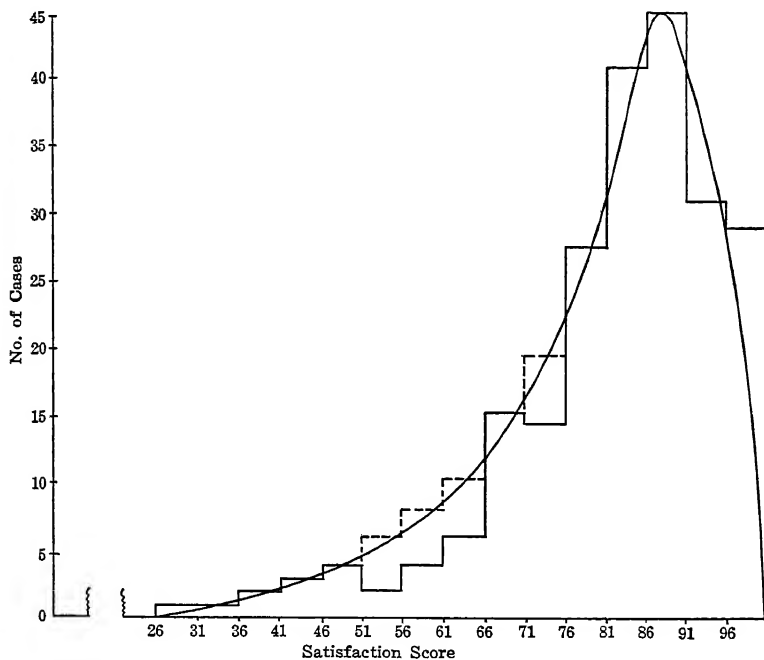
The chief utility of this tension theory is in the field of societal processes. Such processes as "co-operation," "accommodation," "conflict," are defined by behavior relative to desiderata. These and some thirty others can be more precisely defined by equations derived from Eq. 34 or Eq. 49, Ch. V, as described in Chapter X. Even the relationships, the static aspects of these processes, can be better expounded after the dynamic aspects have been developed in the time sector. A critique or attempt to appraise tentatively the evidence bearing on this tension theory seems premature at this writing. Its utility, or lack of it, in the hands of other investigators needs to be explored. Its relations to the highly developed theories of value in Economics need to be worked out. The economic marginal utility theory, the cost of production theory, and others, may shed light which tends to corroborate or refute the  $V$ -theory offered here for a wider field than the phenomena of Economics. But the importance of desiderata for an understanding of societal phenomena is so great, that more research is urgently needed in methods of observing desiderata more objectively and in the theoretical systematizing of such data.

### III. *S-SITUATIONS*

The definitions and formulae of this chapter may be illustrated by the sample of fifteen sets of quantified data which follow. These

have been selected from the fifteen hundred such S-situations in our photostated collection which were the basic facts from which the entire S-theory was induced.

## S. 1



Distribution of marital satisfaction of 221 men and women, excluding clinical cases. The dotted lines indicate marginal cases which were probably selected out.

Ref.: Bernard, Jesse, "Factors in the Distribution of Success in Marriage," *Amer. Jour. Soc.*, Univ. of Chicago Press, Vol. XL, No. 1, July 1934, p. 54.

Descriptive formula:  $S_1 = {}_iI : P_p$   
 Legend:

$S_1$  = The situation  
 records

$I$  = a marital satisfaction score  
 $i$  = in 16 class-intervals  
 with corresponding

$P$  = frequencies of the 221 men and  
 women

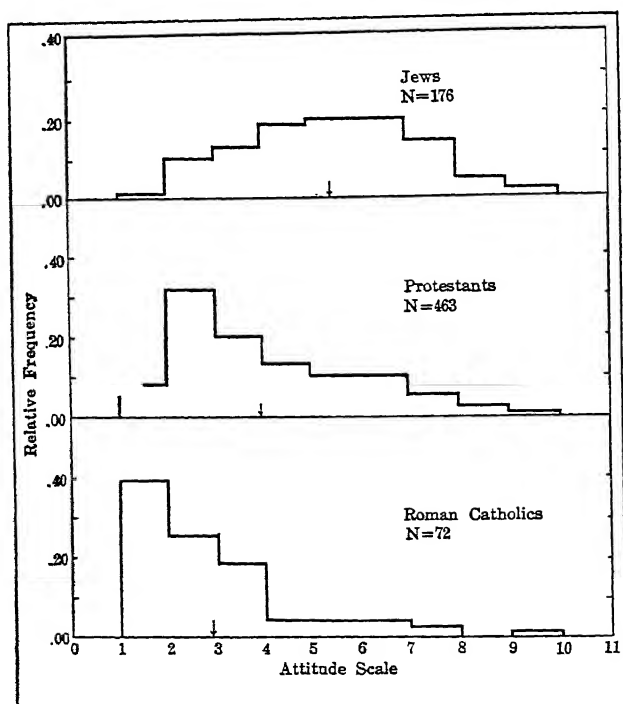
in each of

$|_p = 2$  plurals  $\left\{ \begin{array}{l} \text{actual and} \\ \text{marginally} \\ \text{selected} \end{array} \right.$

*Comment:*

The formula  $\mathcal{I} : \mathcal{P}$  is the standard type for a frequency distribution of people on some characteristic. Note the negative skewness of this particular distribution curve.

## S. 2



Ref.: Thurstone, L. L., and Chave, L. F., *Measurement of Attitudes*, Univ. of Chicago Press, 1929, p. 70.

Descriptive formula:  $S_2 = \mathcal{I} : \mathcal{P}$

Quantic number = 0;1;0;1

Legend:

$S_2$  = The situation

: = with corresponding

records

$\mathcal{P}$  = relative frequency of persons

$\mathcal{I} = 11$  class-intervals of attitude  
towards the church

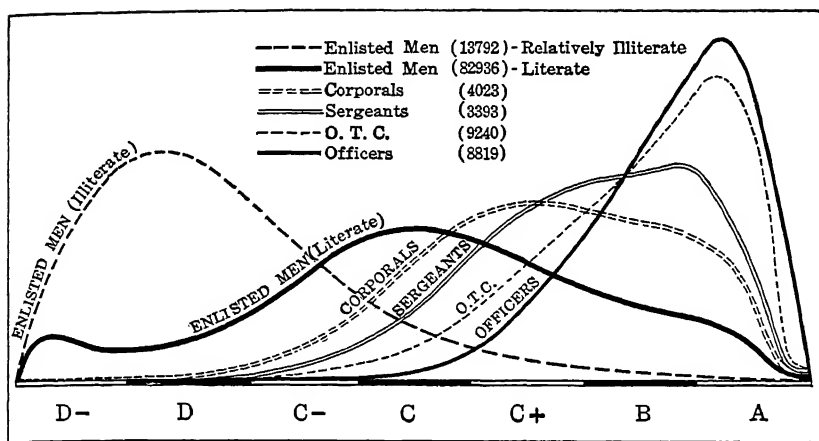
in each of

$\mathcal{P} = 3$  religious plurels

*Comment:*

These three distributions of attitudes towards the church, ranging from most "pro" at the left to most "anti" at the right, show Protestants and Catholics, with positively skewed curves, and means well on the "pro" side of the middle neutrality point.

## S. 3



Ref.: *Army Mental Tests* (Pamphlet), 1918, p. 8.

Descriptive formula:  $S_3 = \underline{i_1}I : P_p$

Quantic number = 0;1;0;1

Legend:

$S_3$  = The situation

records

for each of which there are

$P$  = the number of persons tested

$I$  = an intelligence test indicant

in each of

$\underline{i_1}$  = in 7 major and an unstated number of minor class-intervals

$|_p$  = 6 army rank plurals

Comment on notation:

1. As the size of the class-interval in which the smooth curves are drawn is not stated, their number is underlined ( $\underline{i_1}$ ) to denote indefiniteness, as usual in S-notation.

## S. 4

## COMPARATIVE HYGIENIC STATUS OF BEDOUIN SAMPLES

Sample	N	Mean Score	Standard Deviation
Nomadic Bedouins.....	62	144	13
Partially or recently settled Bedouins.....	140	182	24
Alaouite peasants "Normal" rural sample	100	285	43
Armenian refugees, peasants with urban background .....	32	321	39
Urban sample .....	50	654	60

The differences between means are all statistically highly significant. Note the homogeneity of Bedouin culture as reflected in the very small standard deviation of the nomadic sample.

Ref.: Dodd, Stuart C., *A Controlled Experiment on Rural Hygiene in Syria*, American Press, Beirut, 1934, p. 311.

Descriptive formula:  $S_4 = {}_qP_p : {}^{\mu,\sigma}I$

Quantic number = 0;1;0;1

Legend:

$S_4$  = The situation

records

$|_p = 5$  density plurels  $\left\{ \begin{array}{l} \text{nomadic} \\ \text{to} \\ \text{urban} \end{array} \right.$   
of

${}_q|$  = Arab families

with corresponding

${}^{\mu,\sigma}|$  = mean and sigma  
of

$I$  = an indicant of hygiene

Comment:

The outstanding finding is that hygienic status correlates highly with population density in this situation:

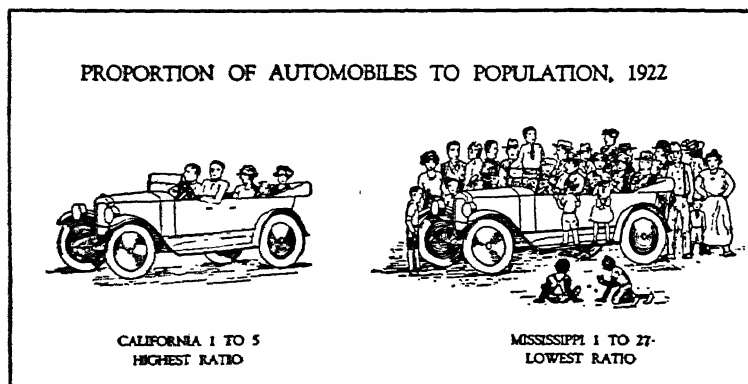
$$I \bullet (PL^{-2}) > 0$$

Comment on notation:

1. The family is the unit of frequency, and is so denoted by the class-interval script,  ${}_q|$ .

2. The use of the point script to specify the mean and sigma points may be noted.

S. 5



Ref.: Rieggleman, R. John, *Graphic Methods for Business Statistics*, McGraw-Hill Book Company, 1936, p. 23.

Descriptive formula:  $S_5 = (PI^{-1})_p$

Quantic number = 0;9;0;1

Legend:

$S_5$  = The situation

records

$P$  = the number of persons

$I^{-1}$  = per automobile  
in each of

$|_p = 2$  states

## S. 6

## FERTILITY OF WOMEN IN DIFFERENT DISTRICTS OF LARGE CITIES

	<i>Paris</i>	<i>Berlin</i>	<i>Vienna</i>	<i>London</i>
Very poor districts.....	108	157	200	147
Poor districts.....	95	129	164	140
Comfortable districts.....	72	114	155	107
Very comfortable districts.....	65	96	153	107
Rich districts.....	53	63	107	87
Very rich districts.....	34	47	71	63

Ref.: Holmes, Samuel J., *The Trend of the Race*, Harcourt, Brace and Co., 1921, p. 132.

Descriptive formula:  $S_6 = {}^iI : (P, P_{,,}^{-1})_p$

Quantic number = 0;1;0;19

Legend:

$S_6$  = The situation

() = fertility index

records

of

${}^iI$  = an ordinal indicant of wealth  
in 6 steps

$P_{,,}$  = births

$P_{,,}^{-1}$  = per woman

for each of which there is a  
corresponding

in each of

$|_p$  = 4 cities

Comment:

As always, the descriptive formula can be written, if preferred, with numbers inserted in the scripts and the legend simplified, thus:

Legend:

$S = {}^6I : (P, P_{,,}^{-1})_4$

$I$  = wealth  $P_{,,}$  = women  
 $P_{,,}$  = births  $|_p$  = cities

For anyone knowing the standardized rules of S-notation (as codified in Appendix II), this legend is fully as informative as the longer one above.

As usual, the descriptive formula can be rearranged so as to state the findings in the data. The outstanding generalizations in this table are: (1) the high negative correlation of fertility rate and wealth in every city, and (2) the consistent differences between cities where Parisian fertility  $(P, P_{,,}^{-1})_p$  is lowest in every wealth class and Viennese  $(P, P_{,,}^{-1})_{p'}$  is highest in every wealth class. These findings stated in symbols are:

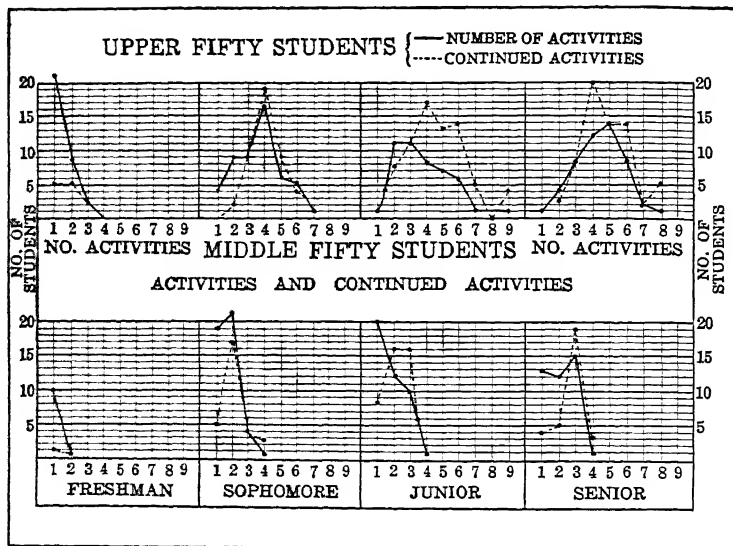
$$1. I \bullet (P, P_{,,}^{-1}) < 0$$

$$2a. {}^iI : (P, P_{,,}^{-1})_{p'} < (p', p'v) < p''$$

or comparing mean fertility rates regardless of wealth:

$$2b. (P, P_{,,}^{-1})_{p''} = p'v + 34 = p'' + 41 = p' + 71 = 142$$

## S. 7



Ref.: Chapin, Stuart F., "Measuring the Volume of Social Stimuli—A Study in Social Psychology," *Social Forces*, Vol. IV, No. 3, Mar. 1926, p. 488.

*Descriptive formula:*  $S_7 = {}_i I_1 : P_p : q$

*Legend:*

$S_7$  = The situation  
records

$I_1$  = 2 indicants of activities

$i_1$  = from 1 to 9 in number

$:$  = with a corresponding

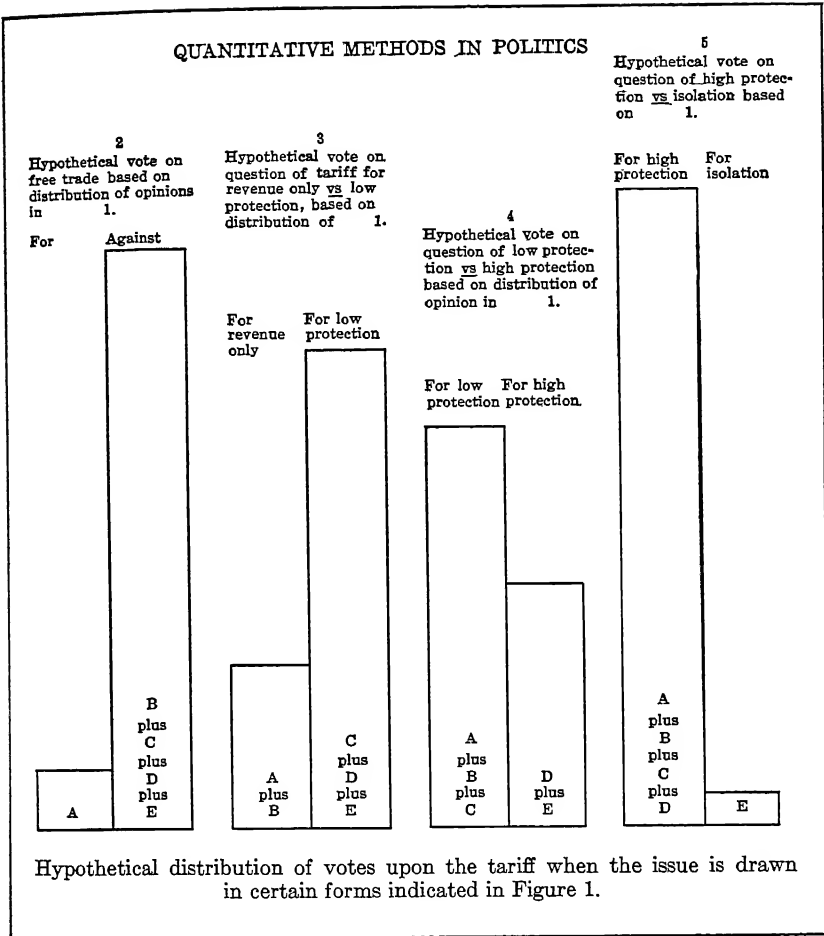
*Quantic number* = 0;1;0;1

$P$  = number of students  
in each of

$p$  = 2 tertile pleurels  
each subdivided into

$q$  = 4 classes (Freshman to Senior)

S. 8



Ref.: Rice, Stuart A., *Quantitative Methods in Politics*, F. S. Crofts & Co., 1928, p. 80.

Descriptive formula:  $S_8 = {}^iI : {}^{z'-i'}, {}^{i'-0}P$

Quantic number = 0;1;0;1

Legend:

$S_8$  = The situation

there is

records

P = a cumulative (unstated) frequency of responders

I = an indicant of tariff opinion

${}^{z'|}$  = from the upper limiting point

${}^i|$  = in 5 ordinal points

and

and

${}^0|$  = from the lower limiting point

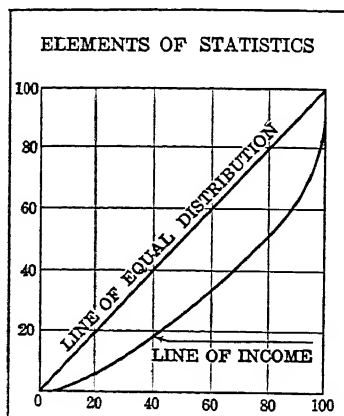
${}^{i'|}$  = for each point (i.e., a "statement")

*Comment on notation:*

1. Note the use of a difference in two point scripts to define the range as the difference between its upper and its lower limit. In this situation,  $S_s$ , each particular point ( $i'$ ) of the five serves in turn as the middle limit, in being simultaneously the lower limit of the range above it ( $i''-i'$ ) and the upper limit of the range below it, ( $i'-i''$ ).

The tops of the columns define two ogive curves (one from lower left to upper right, and another from upper left to lower right), which are of the "more than" and "less than" type of cumulative distribution curves.

## S. 9



Ref.: Davis and Nelson, *Elements of Statistics*, The Principia Press, 1935, p. 28.

Descriptive formula:  $S_s = \%P :: \%I_i$   
Legend:

Quantic number = 0;1;0;1

$S_s$  = The situation is  
a record of

$\%I$  = percentages of total income  
in

$\%P$  = percentages of the total population

$i_i = 2$  classes  $\left\{ \begin{array}{l} \text{"equality"} \\ \text{and} \\ \text{"actuality"} \end{array} \right.$

$::$  = cross-classified  
with

*Comment:*

If income were equally distributed so that its standard deviation were zero, the cumulative percentages of population and income would be equal for any particular percentage and perfectly correlated for all the range.

$\%P = \%I$  or  $\%I \bullet \%P = 1.0$  Line of equal distribution

When the line of actual income is curved, it would be represented by a coefficient of correlation of less than unity.

For income presented in this double percentage form, the correlation coefficient

cient measures the equality of distribution.  $r = 1.00$  at the limit of perfect equality of incomes, and  $r = 0$  at the other limit of maximal inequality, as in perfect monopoly.

## S. 10

## THE FOUR SYSTEMS

We are to be concerned with matrices of the following kind: Each is to contain persons in the row (persons  $A, B, C \dots N$ ) and items such as personality traits, tests, measurements or attributes in general, in the column (attributes  $a, b, c, \dots n$ ).

Persons

		Persons						
		A	B	C	D	...	...	N
Attributes etc.	a	$aX_A$	$aX_B$	$aX_C$	...	...	...	$aX_N$
	b	$bX_A$	$bX_B$	$bX_C$	...	...	...	$bX_N$
	c	$cX_A$	$cX_B$	$cX_C$	...	...	...	$cX_N$
	...	...	...	...	...	...	...	...
	n	$nX_A$	$nX_B$	$nX_C$	...	...	...	$nX_N$

The variate of person A for item a, or of item a for person A, is represented by  $aX_A$ , and the same connotation is held throughout.

Ref.: *Psychometrika*, Vol. 1, No. 3, Sept., 1936, p. 196.

Descriptive formula:  $S_{10} = {}^pP :: I_i$

Quantic number = 0;1;0;1

Legend:

$S_{10}$  = The situation  
records

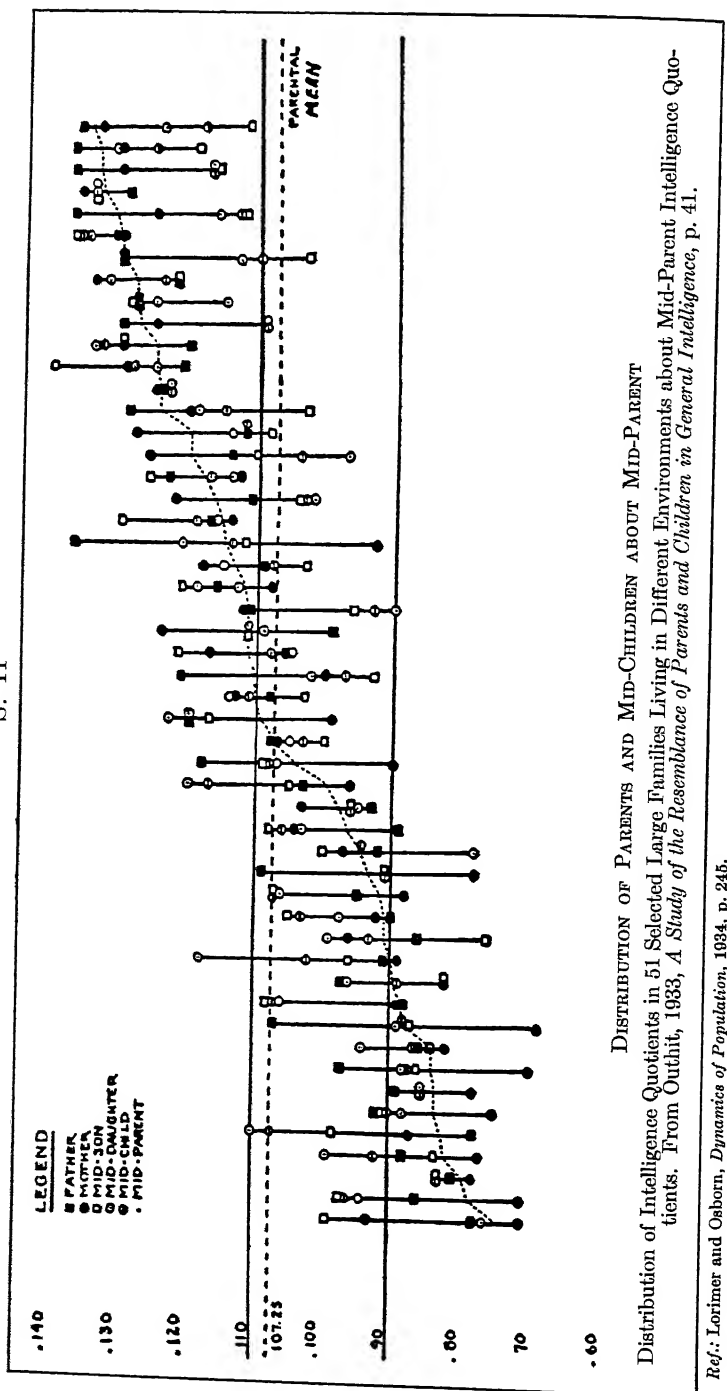
I = an indicant  
of each of

P = N persons  
:: = each having

$|_i$  = n different kinds

Comment:

This second-degree matrix is an aggregation of n "list-distributions." It is a cross-classification of persons and indicants.



S. 11 (*Continued*)

*Descriptive formula:*  $S_{11} = \underline{P}_p : q : r : {}^M I$

*Quantic number* = 0;1;0;1

*Legend:*

$S_{11}$  = The situation

and

records

$|_r$  = 2 sex subplurels

$\underline{P}_p$  = 51 family plurels

with corresponding

subdivided into

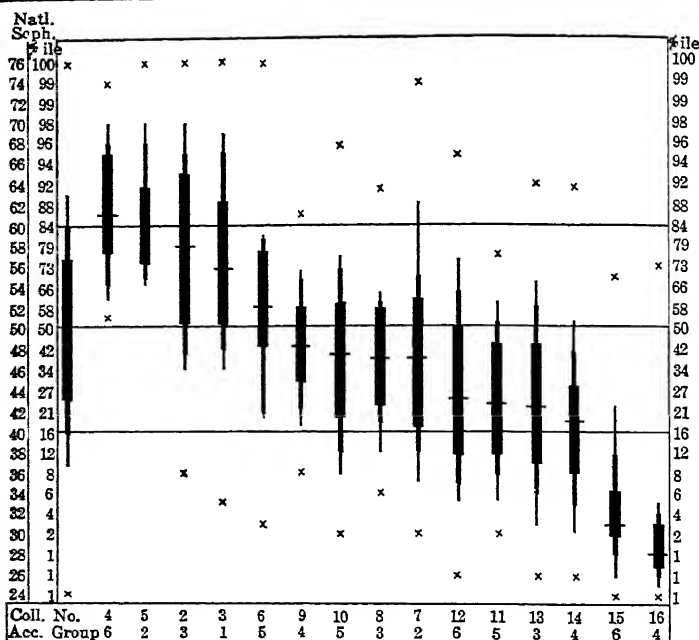
${}^M I$  = mean IQ indicants

$|_q$  = 2 generation plurels

*Comment:*

The versatility of the S-formula with its hierarchical scripts is illustrated here, in describing in algebraic symbols even the most unusual types of graphs or presentations of data.

## S. 12



Variability of achievement in participating colleges as measured by the combined score on the English, literary acquaintance, and general culture tests.

The middle horizontal line shows the national median, and the other two are at the 16th and 84th percentiles of the national sophomore distribution. The first vertical bar represents the national group of 8,537 sophomores from 132 colleges that took this combination of tests, and each of the other bars represents an individual college. The wide portion of each bar represents the range of scores of the middle half of the sophomores in each college. The narrow parts extend to the 16th and 84th percentiles in each college. The lines at the ends extend down to the 10th percentile and up to the 90th percentile. The crosses below the bars represent the lowest scores and those above represent the highest scores in the several colleges. The short cross line at the middle of each bar represents the median score of the college.

While this chart is based entirely on percentiles, the scale has been altered to correspond roughly to a sigma scale, so that vertical distances are approximately comparable. The sigma scale is derived from the percentile scale.

The fifteen colleges included here are chosen to represent the 132 that gave the tests to sophomores. All types of institutions are represented in the fifteen. The numbers of cases vary from 27 to 272. The figures at the bottom of the chart refer to the accreditation groups which are described later in this report.

*Descriptive formula:*  $S_{12} = \underline{P}_p, \Sigma_p : {}^iI$

*Quantic number* = 0;1;0;1

*Legend:*

$S_{12}$  = The situation

each with

records

$I$  = an indicant of achievement

$\underline{P}_p$  = 15 colleges separately

${}^i|$  = with the 5 quartile and 4 other  
percentile points specified

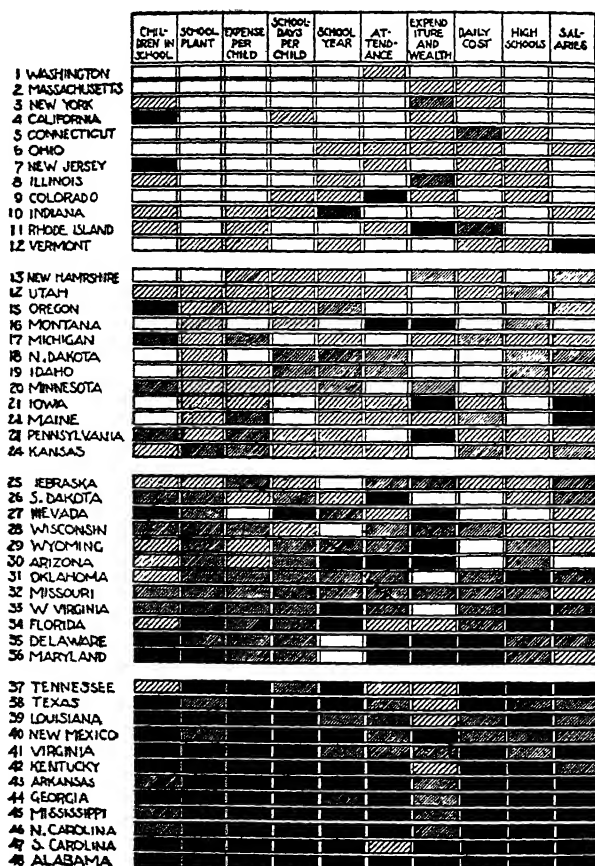
and

$|\Sigma_p$  = together

*Comment:*

This situation is an excellent illustration of the graphing of dispersion so as to compare both different plurels and different measures of dispersion.

## S. 13



Rank of state in each of ten educational features, 1910. White indicates that the state ranks in the highest 12 of the 48, light shading that it ranks in second 12, dark shading that it ranks in third 12, and black that it ranks in lowest 12.

Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, McGraw-Hill Book Company, 1914, page 32.

Descriptive formula:  $S_{13} = \underline{P}_p : T_1$

Legend:

$S_{13}$  = The situation  
records

$\underline{P}_p$  = for the 48 States

Quantic number = 0;1;0;1

on each of

$I_1$  = 10 indicants of education

$\prime$  = its quartile position

*Comment on notation:*

The quartile rank might be alternatively considered as a class-interval and symbolized by ,I, instead of by the point script as here, 'I, denoting a position on an ordinal indicant (ranks). Although there are four ranks, yet there is only one rank corresponding to a particular State and indicant.

The date is omitted for simplicity until discussed in a later chapter.

## S. 14

---

 RANKING OF COTTAGES ACCORDING TO THE SUM OF ATTRAC-  
TIONS AND REPULSIONS IN PERCENTAGES
 

---

<i>Cottages</i>	<i>Attractions</i>	<i>Repulsions</i>
C11.....	85.5%	14.5%
C4.....	77.0%	23.0%
C5.....	77.5%	22.5%
C13.....	74.0%	26.0%
C9.....	69.0%	31.0%
CA.....	69.5%	29.5%
C14.....	67.0%	43.0%
CB.....	65.5%	34.5%
C7.....	65.5%	34.5%
C10.....	66.0%	34.0%
C6.....	58.5%	41.5%
C12.....	58.5%	41.5%
C1.....	58.0%	42.0%
C2.....	52.0%	48.0%
C8.....	47.5%	52.5%

---

Ref.: Moreno, J. L., *Who Shall Survive?* Nervous and Mental Disease Publishing Co., 1934, p. 101.

*Descriptive formula:*  $S_{14} = \underline{P}_p : \%I_i$

*Quantic number* = 0;1;0;1

*Legend:*

$S_{14}$  = The situation

$\%I$  = percentage indicant of attitude

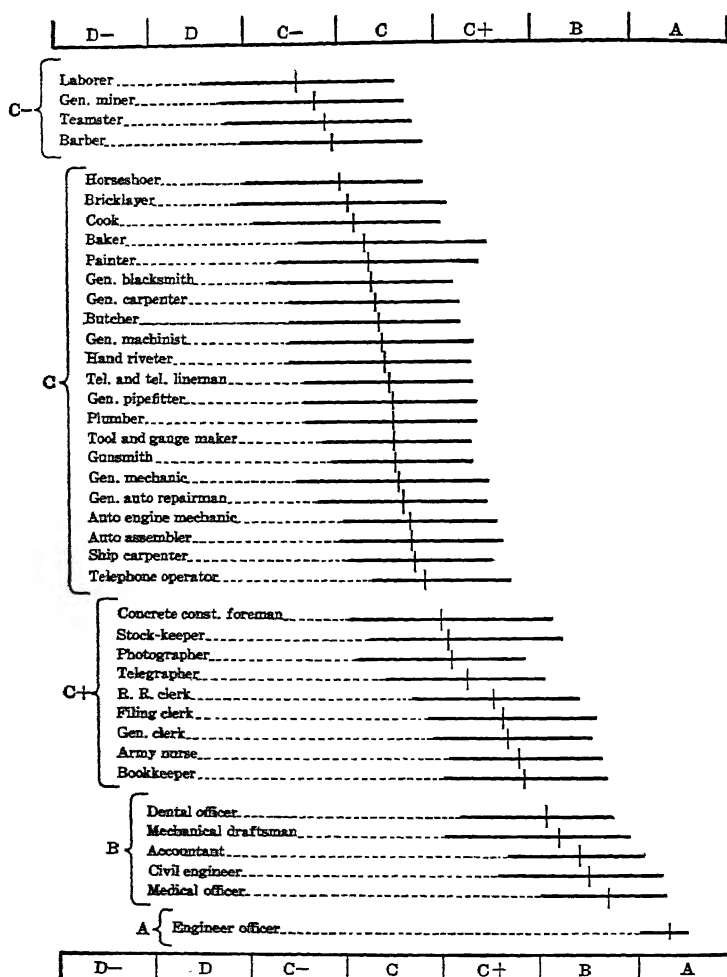
is a record of

$|_i$  = of 2 kinds (attraction, repul-  
sion)

$\underline{P}_p$  = 16 cottage plurels

with a corresponding

## S. 15



Mental test rating of 18,423 men by occupational groups. The length of the bar indicates the range of the middle fifty percent of each group. The cross-bars indicate the median scores. How much of such social stratification is due to inherited mental level and how much to environmental factors is one of the most profound problems of the age. (From R. M. Yerkes, "Psychological Examining in the U. S. Army," *Memoirs*, National Academy of Sciences, Vol. XV, 1921.)

*Descriptive formula:*  $S_{15} = {}_1(P_p : {}^2I)$

*Quantic number* = 0;1;0;1

*Legend:*

$S_{15}$  = The situation

and these grouped into

is a record of

$|$  = 5 letter grades of intelligence

$P$  = 18,423 persons in U. S. army

with their corresponding

grouped into

${}^2I$  = 3 quartile-points of mental test

$|_p$  = 40 occupational plurels

rating

*Comment on notation:*

The fact that the class-intervals of the indicant also control the grouping of the occupational plurels is suggested by writing the  $|$  outside the parenthesis, so that it operates upon the  $P$  as well as upon the  $I$ .

### III. NOTES

1. A societal distribution of a characteristic is thus defined by the quantic formula:

$I^{+1}; P^{+1}$  = quantic formula of a societal distribution of a characteristic  
(Eq. 2, Ch. V)

${}_pP : I$  = a list-distribution of persons (Eq. 1b, Ch. V)

An example of such a list of scores is given in algebraic symbols in each row of S. 10, Ch. V. The list may be of the indicants of an aggregation of plurels, as in S. 4, 12, 13, 14, 15, Ch. V, and the formula for this is:

$\underline{P}_p : I$  = a list-distribution of plurels (Eq. 1c, Ch. V)

${}_iI : \underline{P}_p$  = a frequency distribution of plurels (Eq. 1d, Ch. V)

(See S. 1, 2, 7, 8, Ch. V, for examples.)

Note that in the list-distribution the population index dominates, i.e., it is the independent variable for each of whose values the corresponding (statistically dependent) variable is recorded. In the frequency distribution the indicant dominates (as symbolized by its preceding the colon) and the population index is dependent (as shown by its following the colon).

The societal frequency distributions are defined by having persons or plurels as the units of frequency. Of course, in biological and physical fields distributions may have indicants as units of frequency. Also, the distributions above are of characteristics (i.e., not of time and space). Temporal distributions of population, as in a population pyramid, and spatial distributions are classified by their quantic formula into Chapters VIII and IX respectively. The general formula for a frequency distribution, whether societal or not, whether in the indicatory sector or not is:

${}_i(I), : I,,$  = general formula for a frequency distribution (Eq. 3, Ch. V)

in which the index ( $I$ ) is any function of time, space, population, or indicants. In the frequency distribution, the units of the independent index preceding the colon, may be termed the "distributing units," and the corresponding units of

the dependent index which follows the colon may be termed the "distributed units," or units of frequency.

A special case occurs in distributions of plurals-by-size, where the distributing and distributed units are both parties. The formula here is:

$${}_p:\Sigma_a P = \text{a distribution of plurals-by-size} \quad (\text{Eq. 4, Ch. V})$$

This is a condensation for  ${}_p P : \Sigma_a P$ , which states that for every class-interval of persons there is a corresponding number of plurals. For example in S. 11, Ch. IV, the plurals are families,  ${}_a|$ , specified in size as having one to four wives,  ${}_p P$ . The condensation follows the usual notational principle in S-theory, of expressing all primary indices of one sector by means of descripts attached to the single capital base letter which denotes the type of index.

In terms of matrices, a distribution is one array. It is a first-degree matrix as shown by its one multiple aggregative descript. It is of order  $1 \times |p|$  for the list of persons type, of order  $1 \times |{}_p|$  for the list of plurals type, and of order  $|i| \times 1$  for the frequency types. Thus, in S. 10, Ch. V, each row is a list-distribution of  $N$  persons ( $|p| = N$  in this S). The row is an aggregation of  $N$  persons whose scores are the  $N$  cell entries. The whole second-degree matrix of this S. 10, Ch. V is an aggregation of  $n$  list-distributions of the  $n$  different indicants. (In S-notation,  $|i| = n$ , in this graph.)

Vectorially, an array of a matrix is termed an array-vector. Any distribution then may have as its vectorial formula, in S-notation:

$$\bar{I}P ? = \text{vectorial formula of a distribution} \quad (\text{Eq. 5, Ch. V})$$

By this hypothesis, the indicant is the vector whose direction is given by the implicit attribute (see Eq. 9, Ch. III), and the  $P$  is a pure population whose plurals are defined by the class-intervals of the indicant and are scalar units or coefficients weighting the class-interval units of the indicant vector. This hypothesis is untested and its utility, if any, is unknown. It is recorded here simply to include a possible geometric interpretation for all of the S-theory.

From another geometric point of view the distributed units, or frequencies of a distribution can be represented as units of *area* under the curve of that distribution, while the distributing units measure the abscissa. For a given type of distribution curve, such as the normal curve, the area between given ordinates (i.e., between specified points on the distribution indicant) is the *probability* of the occurrence of parties in that interval. When the whole area under the curve (i.e., the total population) is taken as unity, any subarea between ordinates is expressed as a proportion of unity which is the coefficient of probability.

2. A moment of any order about the mean is defined by:

$$\mu_n = \frac{\Sigma_d I^i}{P} \quad (\text{Eq. 6, Ch. V})$$

For a model of a weighing device to calculate the first and second moments by using the physical angular moments or moments of force, see S. 20, Ch. VI. For the formulae for the statistical moments, their standard errors, and their commoner functions, see Ref. 22, pp. 110-112.

3. A negative exponent as on the P does not affect the dimensionality. It is ignored in reckoning the dimensionality. It denotes a divisor which merely changes the size of the units in which the dividend is expressed. Vectorially, since division by a vector is ambiguous and debarred in vectorial algebra, any index with a negative exponent must be considered as a scalar quantity, a mere coefficient.

4. A general formula for the weighted mean in a population of persons, in S-notation, is:

$${}_M I = \left( \frac{{}^1 W {}^1 I^e + {}^2 W {}^2 I^e \dots + {}^P W {}^P I^e}{{}^1 W + {}^2 W + \dots + {}^P W} \right)^{1/e} \quad (\text{Eq. 12a, Ch. V})$$

When the weight indicants, W, are unities this simplifies to:

$${}_M I = \left( \frac{\sum I^e}{P} \right)^{1/e} \quad \begin{array}{l} (\text{Eq. 12b, Ch. V}) \\ (\text{Ref. 58, p. 6, 7}) \end{array}$$

which when:

$e = -\infty$   ${}_M I = {}^A I$ , the minimal I or lowest limit of the indicant  
(Eq. 12c, Ch. V)

$e = -1$   ${}_M I$  = the harmonic mean (Eq. 12d, Ch. V)

$e = 0$   ${}_M I$  = the geometric mean (upon evaluating indeterminate quantities)  
(Eq. 12e, Ch. V)

$e = +1$   ${}_M I$  = the arithmetic mean (cf. Eq. 8, Ch. V) (Eq. 12f, Ch. V)

$e = +2$   ${}_M I$  = the root mean square,  $= \sigma$  if  $I = {}^{1-M} I$  (i.e., if I is in units of deviation from the mean) (Eq. 12g, Ch. V)

$e = +3$   ${}_M I = Sk_2$  a mean measure of skewness (provided I is in sigma units)  
(Eq. 12h, Ch. V)

$e = +\infty$   ${}_M I = {}^Z I$  the maximal or upper limit of the indicant (Eq. 12i, Ch. V)

The standard error of the arithmetic mean, M, due to random sampling is:

$$\sigma M = \sigma I P^{-.5} \quad (\text{Eq. 13, Ch. V})$$

For other measures of central tendency, such as the median and the mode, and for the properties and computation of all the indices of distributions, the student should study a standard textbook in Statistics.

5. The equation of this curve is:

$$y = .3989 / (2.7182)^{.5} \sigma^{-1} = \text{the normal probability curve in terms of unit-area and unit-sigma, i.e., } P = I \text{ and } \sigma = 1$$

(Eq. 15, Ch. V)

In this equation y is the ordinate,  ${}_M I$ , the abscissa in sigma units. The areas under the curve between any two ordinates are proportions of the population. These quantities are tabulated in the table of the normal probability integral so that for given amounts of a characteristic the corresponding proportion of the population, or probability for a person, may be read off.

6. The student beginning quantitative sociology should distinguish the connotations of the terms "standard deviation" and "standard error." The former is usually an observed quantity, calculated from the data of a given distribution

by Eq. 9, Ch. V. The latter is a quantity calculated by a formula which is specific to the index of which it is the standard error. The scripts attached to a sigma show whether it is a standard deviation of a distribution of observed data, or a theoretic standard error of sampling of an index. For example the standard errors of Eqs. 8-11, Ch. V, are:

- a.  $\sigma_M = \sigma_{IP}^{-.5}$  = standard error of an arithmetic mean  
(Eq. 16a, Ch. V)
- b.  $\sigma_\sigma = \sigma_{I(2P)}^{-.5}$  = standard error of a standard deviation of a normal distribution  
(Eq. 16b, Ch. V)
- c.  $\sigma_{Sk_1} = 2.4495P^{-.5}$  = the standard error of an index of skewness  
(Eq. 16c, Ch. V)
- d.  $\sigma_{\beta_2} = 4.899P^{-.5}$  = the standard error of an index of kurtosis  
(Eq. 16d, Ch. V)

Note that the standard error, measuring unreliability of any index, varies inversely with the square root of the population,  $\sigma(I) \propto 1/P^{.5}$ . Thus, to cut a standard error in half, the population observed must be quadrupled; to cut it to one third requires nine times the original population, etc.

7. The statistical significance of any difference is the ratio of the difference to its standard error. If this ratio is three or greater it is conventionally called significant, as the probability table shows that it would occur by chance about once in a thousand samples. In the example above of a difference between means of plurals:

If  $\sigma(M_{I_{p'-p''}}) > 3$  the difference is significant (Eq. 17, Ch. V)

(In cross-script notation,  $M_{I_{p'-p''}}$  states the difference in mean indicants  $M_I$  of plurals  $|_{p'}$  and  $|_{p''}$ )

The standard error of such a difference in mean indicants is:

$$\sigma(M_{I_{p'-p''}}) = (\sigma_{p'}^2 + \sigma_{p''}^2)^{.5} P^{-.5} \quad (\text{Eq. 18a, Ch. V})$$

The general formula for the standard error of a difference in indicants which are not means and which may be correlated is:

$$\sigma_{(i,j)} = (\sigma_i^2 + \sigma_j^2 - 2r_{ij}\sigma_i\sigma_j)^{.5} \quad (\text{Eq. 18b, Ch. V})$$

8. See S. 3, Ch. XII. Note that the correlation coefficient is a product-moment reduced to sigma units of the variables correlated,  $\frac{\sum IJ}{P \sigma_I \sigma_J} = \frac{\sum \sigma_I \sigma_J}{P} = r$ . As the coefficient of determination,  $r^2$ , and the coefficient of nondetermination,  $k^2$ , add up by definition to 1, they are more correctly interpreted in percentage terms than is legitimate for the coefficient of correlation ( $r$ ) or the coefficient of alienation ( $k$ ).

9. The standard S-notation for percentiles is:

- $_{p'}:_{p'}P$  = the number of persons in each of  $p$  equal class-intervals to each of which correspond a particular class-interval of the indicant
  - $_{p'}:_{p'}I$  = an indicant with class-intervals corresponding to the  $p$  equal plurals.
- For percentiles,  $_{p'}| = 100$ ; for deciles,  $_{p'}| = 10$ ; for quartiles,  $_{p'}| = 4$ , etc.

10. The formula for monopolistic dispersion is:

$$\max. \sigma_v = \Sigma I_v P^{-.5} = {}^M I_v P^{-.5} \quad \text{Dispersion in a monopoly} \quad (\text{Eq. 19a, Ch. V})$$

i.e., the situation of a maximal sigma.  ${}^M I_v$  is here the mean amount of the indicant held by the population  $P$ , times the square root of  $P$ . The subscript  $v$  identifies the indicant as a value-indicant, i.e., an indicant of some desideratum, since usually, monopoly is of societal interest only for things which are desired

The derivation of this sigma of a monopoly is:

If, as usual,  ${}^v I =$  the quantity of the indicant possessed by one person

$$\Sigma I = \sum_1^P I = \text{the total quantity possessed by all persons}$$

$${}^M I = \Sigma I P^{-1} = \text{the average quantity possessed by the population, } P$$

The standard deviation using the formula for an arbitrary origin is:

$$\sigma_v = \left[ \sum_1^P ({}^v I)^2 P^{-1} - ({}^M I)^2 \right]^{.5} = \text{sigma of the indicant} \quad (\text{Eq. 19b, Ch. V})$$

Now in monopoly one person's  ${}^v I = \Sigma I$  and for every other person  ${}^v I = 0$ , hence  $\Sigma ({}^v I)^2 P^{-1} = (\Sigma I)^2 P^{-1}$ , and Eq. 19b becomes:

$$\sigma_v = ((\Sigma I)^2 P^{-1} - (\Sigma I^2 P^{-2}))^{.5} = \Sigma I P^{-1} (P - 1)^{.5} = {}^M I (P - 1)^{.5}$$

Accurate sigma of a monopoly (Eq. 19c, Ch. V)

Whenever  $P$  is large enough so that the difference between it and  $P - 1$  is negligible, we may write  $\sigma_v = {}^M I P^{.5}$ , which is Eq. 19a. For  $P = 50$ , the error of the approximative formula, Eq. 19a, is 1%; for  $P = 25$ , the error is 2%.

11. For examples of dissimilarity see S. 19, 23, 35, Ch. X.

For examples of dispersion see S. 44, Ch. X; S. 21, 26, Ch. XII.

For examples of variation see S. 13, 15, Ch. V; S. 34, Ch. VI; S. 25, Ch. X; S. 66, Ch. X (definitions vague).

12. Dudycha seems to identify the normal curve with a mesokurtic one. His conclusions assert normality, which he tests nowhere (see Eq. 22a, Ch. V), when they should assert mesokurtosis only. It may also be noted that, even with mesokurtosis established, Allport's assertion may still be true that the double J-curves of non-telic conforming behavior are positively accelerated in both slopes. Best fitting curves and their goodness-of-fit coefficient would be a further useful research upon these distributions.

Dudycha tested for kurtosis by the interpercentile formula which is briefer, but usually less rigorous, than using the fourth moment as in Eqs. 11 and 16d, Ch. V. Kelley's interpercentile formula (Ref. 35, p. 77), which he uses, is:

$$Ku = Q/D \text{ the interpercentile measure of kurtosis} \quad (\text{Eq. 22a, Ch. V})$$

where  $Q =$  the quartile and  $D =$  the 10 to 90 percentile range

$${}^c Ku = .27779 P^{-.5} \quad (\text{Eq. 22b, Ch. V})$$

where  $Ku = .26315 =$  mesokurtosis (Eq. 22c, Ch. V)

$Ku > .26315 =$  platykurtosis (Eq. 22d, Ch. V)

$Ku < .26315 =$  leptokurtosis (Eq. 22e, Ch. V)

Allport's test of conformity was that 50% of the population (or of the instances observed) should fall in the extreme or completely conforming class-interval (as in Catholics genuflecting "completely" in church vs. kneeling "halfway" vs. not kneeling). Dudycha refines this to equate the last term of a geometric series to half the population, shows that a geometric ratio,  $r$ , of 2 is the boundary of "conforming," and also defines Allport's J-curve. Both authors seem to have neglected the vital effect of the number of class-intervals, if their criteria are to be generalized, for data in three class-intervals might have such a frequency in the "conforming" class-interval as to pass their criteria of conformity, but when expressed in six class-intervals it might have less than half the chance of passing the criteria. To allow for this, the criterion should be stated in terms, not of the extreme class-interval, but of some proportion of the range, perhaps  $1/5$ , at the conforming extreme. With this correction Dudycha's formula in S-notation is:

$(1-r)^{1/5} \cdot I^4 = {}_rP =$  a criterion of conformity defining Allport's J-curve hypothesis in terms of a geometric series of 5 class-intervals

(Eq. 23, Ch. V)

In this, the particular extreme class-interval,  $|$ , of the indicant  $I$ , (Allport's "telic continuum"), which is defined by one fifth of the interval between the points that limit its range,  $'-1|$ , and multiplied by the geometric ratio,  $r$ , of the series of 5 class-intervals, is equated to a percentage of the population,  ${}_rP$ . Instead of Allport's fixed point ( $.5P$ ) making the dichotomy of "conforming" or "non-conforming," it would be better to express the degree of conformity by stating the observed numerical coefficient of  $P$  as a "percentage of conformity,"  ${}_rP$ . Allport's criterion  ${}_rP = 50\%$  means "a majority conforming."

This percentage of conformity criterion is in line with modern non-Aristotelian logic in science. The Aristotelian law of identity, that all is  $A$  or not- $A$ , is being found an increasingly inadequate approximation to truth. There are degrees of identicalness. (See Ref. 36.) Probability and correlation coefficients and numerous other measures express degrees of similarity with identity as a limit. The facile dichotomizing of truth into "true" and "not true" should be outgrown in scientific thinking.

Social scientists need to get beyond the primitive thought habits of all-or-none indicants and think in terms of calibrated cardinal indices as far as these can be developed. When this is done, the conventional defining of a point on such an index to set boundaries between class-intervals, as Allport's criterion aimed to do, is both desirable and necessary.

*Note added to galley proof January 1940:*

The comment above on Allport's hypothesis was penned along with the rest of Chapter IV some three and a half years before this volume will reach publication in print.

Subsequent research reported in the sociological literature has developed the hypothesis much further. This comment is not deleted here, however, as it serves both to illustrate S-theory notation and the primitive dichotomizing habits of many social scientists.

13. These types of distribution curves have been charted by Karl Pearson and others. For an excellent résumé of the chief types, their equations, properties, and identifying criteria see Ref. 35, Ch. VII.

For example, the criteria identifying the simplest types which are most frequently used in sociology are:

$\beta_1 = 0, \beta_2 = 3 =$  criteria for the normal curve (Eq. 24a, Ch. V)

$\beta_1 = 0, \beta_2 = 1.8 =$  criteria for the rectangular distribution (ranks, <sup>1</sup>I)

(Eq. 24b, Ch. V)

$\beta_2 - \beta_1 - 1 = 0 =$  criteria for the two-category distribution (all-or-none, <sup>2</sup>I)

(Eq. 24c, Ch. V)

$\beta_1 = 0, \beta_2 = 1$  if the two categories are of equal frequency (Eq. 24d, Ch. V)

$\beta_1 = 0$  is the criterion of all symmetric curves as in asymmetric curves  $\beta_1 \neq 0$

(Eq. 24e, Ch. V)

The general procedure to find the type and equation for given data is to calculate  $\beta_1$  and  $\beta_2$  by Eqs. 10 and 11, Ch. V, and find the corresponding point or line in Chart XVIII, p. 127, of Ref. 35, with due regard for approximateness shown by the appropriate standard errors.

14. This hypothesis that evolution is a trend towards the stable has an interesting parallel in the physical principle of entropy expressed in the second law of thermodynamics, that free energy in the universe tends to become less and less. It becomes cumulatively bound or made unavailable in the equilibria of closed systems which are stable. Is normality of distribution a biological working out of this principle of entropy? Here lies a field of research between Physics and Biometrics.

15. For an amplification of the scientific definition and measurement of desiderata see Ref. 14.

16. The above definition of minimality and list of the chief types may be illustrated by more than one fourth of the S-situations presented in this volume. These are tabulated below in a list giving the general characteristics socially valued and the index measuring it. As none of these characteristics have had their value experimentally determined by measuring their desirability to the plurel of these situations in which they are a social problem, their identification is somewhat subjective. There would be general agreement that malaria is undesirable, but views on polygyny, armaments, and anti-free-speech laws would probably be more controversial in some plurels. Wherever there is greatest uncertainty as to the percent of people to whom a characteristic would be a positive or a negative value, it is so indicated by a "7" in the footnote column.

About 87% of the 88 situations in this sample involve a population explicitly. About one eighth of the situations involve the characteristics of a social problem but have so remote a reference to a social problem *plurel* as to be written with the population as nul (i.e., with a fourth quantic digit of 0). These are marked "8" in the footnote column.

## EXAMPLES OF MINIMALITY

*in the 300 S-situations of this volume*

<i>S-situation No. and Chapter</i>	<i>Positively or Negatively Valued Characteristics</i>	<i>See Foot- notes</i>	<i>Index Measuring the Characteristic</i>	<i>Quantic Number</i>
<i>Domestic:</i>				
S. 3, Ch. II	Divorce	3	Marriages/divorce	8;0;0;19
S. 3, Ch. IV	Social security	6	Plurels aided	0;0;0;1
S. 11, Ch. IV	Polygyny	7	% families	0;0;0;19
S. 1, Ch. V	Marital satisfaction	1	Scores	0;1;0;1
S. 11, Ch. X	Miscel. deficiencies	2, 6	Families by types	9;0;0;1
S. 43, Ch. XI	Falling birth rates	2, 3, 7	Confinements/10 fam- ilies	81;0;0;1
S. 5, Ch. XII	Marital adjustment	1	Scores and attributes	0;2;0;1
<i>Medical (and vital):</i>				
S. 18, Ch. II	Safety	1	Deaths/passenger miles	9;0;9;1
S. 30, Ch. II	Hygiene	1, 3	Scale scores	9;1;0;1
S. 12, Ch. III	Sanitation	6, 8	Scale scores	0;1;0;0
S. 3, Ch. IV	Social security	6	Plurels aided	0;0;0;1
S. 9, Ch. IV	Defects	1	% persons by types	0;0;0;1
S. 6, Ch. X	Defects and mortality	1	% persons by ages	9;0;0;1
S. 7, Ch. X	Mortality	2	Population by causes	9;0;0;1
S. 9, Ch. X	Injuries	2	% by causes and sex	9;0;0;1
S. 12, Ch. X	Mortality	1	% by countries	9;0;0;1
S. 34, Ch. X	Infant mortality	2	%	9;1;0;1
S. 43, Ch. X	Sickness cost	1	% families, % cost	9;1;0;1
S. 62, Ch. X	Yellow fever	3, 4, 8	Area	9;0;2;0
S. 82, Ch. X	Infant mortality	2	% by classes	19;0;0;1
S. 85, Ch. X	Sickness of pupils	2	No. and % absent	2;0;0;1
S. 19, Ch. XI	Malaria	5	Doctor's visits	8;1;0;1
S. 7, Ch. XI	Diphtheria	5	Cases	8;0;0;1
S. 42, Ch. XI	Mortality	2, 3	Deaths/pop. by ages	81;0;0;1
S. 13, Ch. XII	Mortality	1, 2	Deaths/pop. by ages and diseases	19;0;0;1
S. 11, Ch. XII	Illness by income	1, 2	Cases/pop. by ages	6;0;0;1
S. 18, Ch. XII	Insanity	3	Index number	9;0;0;1
<i>Economics:</i>				
S. 4, Ch. II	Unemployment	2	Weeks of duration	2;0;0;1
S. 32, Ch. II	Standards of living	6	Pop. and production %	8;1;0;1
S. 3, Ch. IV	Social security	6	Plurels aided	0;0;0;1
S. 4, Ch. IV	Waste	1	Man power	0;0;0;1
S. 8, Ch. IV	Technological displacement	2	Workers displaced	0;0;0;1
S. 9, Ch. V	Unequal income	6, 7	% pop. and % in- come	0;1;0;1
S. 27, Ch. VI	Unemployment	2	%	0;2;0;1
S. 14, Ch. X	Unemployment	3	Population	9;0;0;1
S. 36, Ch. X	Distribution of income	1	% families, % income	9;1;0;1
S. 43, Ch. X	Sickness cost	1	% families, % cost	9;1;0;1
S. 30, Ch. XI	Cost of social services	3	Dollars per capita	8;1;0;9
S. 32, Ch. XI	Costs of living	2, 3	Index numbers	8;1;0;1
S. 20, Ch. XII	Distribution of income	1	Dollars	9;1;0;1

## EXAMPLES OF MINIMALITY—(Continued)

*in the 300 S-situations of this volume*

<i>S-situation No. and Chapter</i>	<i>Positively or Negatively Valued Characteristics</i>	<i>See Foot- notes</i>	<i>Index Measuring the Characteristic</i>	<i>Quantic Number</i>
<i>Economics:—(Continued)</i>				
S. 7, Ch. XII	Poverty cycle	2, 3	Sigma	1;1;0;1
S. 18, Ch. XII	Pauperism	3	Indices	9;0;0;1
<i>Political (and legal):</i>				
S. 13, Ch. II	Armament	6, 7, 8	Dollars for education	0;1;0;0
S. 13, Ch. III	Armament	6, 7, 8	Dollars for farming	0;1;0;0
S. 2, Ch. IV	War casualties	2	Persons	0;0;0;1
S. 7, Ch. VII	Underworld government	6	Attributes of names	0;0;0;2
S. 17, Ch. VIII	Free speech	4, 7	Attributes of legisla- tion	0;1;2;0
S. 50, Ch. X	Murder	2	% by sex and race	9;0;0;2
S. 24, Ch. X	War	6, 8	Dollars	9;1;0;0
S. 28, Ch. X	Crime	6, 8	Index	9;1;0;0
S. 63, Ch. X	Lynching	1, 4	Persons by States	9;0;2;1
S. 74, Ch. X	Crime	2	Arrests by age	19;0;0;1
S. 82, Ch. X	Legitimacy	2	Infant mortality by classes	19;0;0;1
S. 15, Ch. XII	War	3	War years/century	91;0;0;1
S. 18, Ch. XII	Delinquency	3	Index numbers	9;0;0;1
S. 25, Ch. XII	Robbery	3	Cases	8;1;0;1
<i>Racial:</i>				
S. 2, Ch. VII	Prejudices	7, 2	Ordinal attitude in- dicant	0;1;0;2
S. 1, Ch. VII	Prejudices	1, 7	Cardinal social dis- tances	0;1;0;2
S. 12, Ch. VII	Prejudices	1, 7	Ordinal social dis- tances	0;1;0;11
S. 50, Ch. X	Attitudes	2	Murders by sex and race	9;0;0;2
S. 63, Ch. X	Lynching	1, 4	Persons by States	9;0;2;1
<i>Educational:</i>				
S. 13, Ch. V	Features of school systems	1	Rank by States	0;1;0;1
S. 12, Ch. V	Achievement (academic systems)	1	Scores by colleges	0;1;0;1
S. 29, Ch. VI	Abilities of the feeble- minded	2	Test scores	0;2;0;1
S. 30, Ch. VI	Failures in college	1	Scores and %'s	0;2;0;1
S. 32, Ch. VI	Failures in college	1	Scores and %'s	0;2;0;1
S. 9, Ch. IX	Truancy	2, 3	Days absent	1;1;0;1
S. 70, Ch. X	Intellectual development	1, 3, 4	Score	9;1;2;1
S. 85, Ch. X	Absences	2	% Sick	2;0;0;1
S. 4, Ch. XII	Failure in college	1	Rank and %'s	0;2;0;1
S. 21, Ch. XII	Abilities of the feeble- minded	1	Test scores	9;1;0;1
<i>Recreational:</i>				
S. 2, Ch. II	Leisure	3	Hours of work	91;0;0;1

## EXAMPLES OF MINIMALITY—(Continued)

in the 300 S-situations of this volume

<i>S-situation No. and Chapter</i>	<i>Positively or Negatively Valued Characteristics</i>	<i>See Foot- notes</i>	<i>Index Measuring the Characteristic</i>	<i>Quantic Number</i>
<i>Esthetic:</i>				
S. 1, Ch. III	Noise	6, 8	Attributes	0;0;0;0
S. 18, Ch. VIII	Soot	4, 8	% deposited	0;1;2;0
<i>Ethico-religious:</i>				
S. 12, Ch. II	Interreligious attitudes	1	Cardinal social dis- tance scores	0;1;0;2
S. 24, Ch. II	Motivation in work	1	Ordinal types	9;1;0;2
S. 4, Ch. IV	Vice	1	Persons involved	0;0;0;1
S. 10, Ch. IX	Honesty	1	Scores from tests	1;1;0;1
S. 18, Ch. X	Beliefs	1, 3	% endorsers	9;0;0;1
<i>Linguistic:</i>				
S. 2, Ch. VIII	Diversity of languages	4	Geographic areas	0;0;2;0
<i>Communal:</i>				
S. 21, Ch. VIII	Civilization	1, 4, 8	Ordinals by regions	0;2;2;0
<i>Densitive:</i>				
S. 12, Ch. VIII	National density	1, 4	Pop./square mile	0;0;2;1
S. 14, Ch. VIII	National agricultural density	1, 4	Pop./arable square mile	0;0;8;1
S. 10, Ch. VIII	Continental density	6, 4	Population	0;0;2;1
S. 15, Ch. VIII	Housing density	2	Persons/room	0;0;8;1

*Key to digits in the "Footnote" column*

- 1 = situations presenting a distribution of a valued index in which the lowest negative deviates are the "minimals," the social problem plural as we have defined it
- 2 = situations presenting characteristics of a population of minimals
- 3 = situations presenting a change of characteristics of a population of minimals
- 4 = situations presenting regional distribution of a population of minimals
- 5 = situations presenting therapy of a population of minimals
- 6 = situations presenting some discussion of a population of minimals
- 7 = situations presenting doubtful evaluation, depending on the plural which evaluates
- 8 = situations presenting nul populations, i.e., emphasizing the valued characteristic rather than the plural characterized

17. A deeper issue is here involved. Since people participate in, as well as observe, societal phenomena, "discovery" (connoting observation by a third party of something already existing) may be less accurate a term than "creation," connoting an interaction between research invention (or something previously non-existent in that particular form) and its application (which brings it into existence in human experience just as truly as Columbus' discovery of a new world).

This hypothesis of a natural range (or more accurately, a normal range) illustrates one concept of "natural law" in the social sciences. In the physical sciences man, as an outside observer, notes and states the uniformities (or lack of them, or degrees of them) that exist under specified conditions which he usually manipulates in experimental set-ups. In the social sciences, as a participant

element in the phenomena observed, man may passively observe existing uniformities, but he may also create them. Cultural phenomena, such as the distribution of wealth, may not be entirely normally distributed. (In the United States when incomes are stated in log units, a normal curve has been found to fit well up to the top 15% or so (Ref. 57, p. 522).) But if a normal distribution were adopted as a value, i.e., a standard (because of its biological frequency of occurrence and reasonableness on many grounds), steps to control deviations from it could be taken by society and the law of the normal curve would tend progressively to "hold good," i.e., to describe the existing phenomena. System and order in the social sciences and the consequent prediction and control of human welfare are more likely to be achieved than to be discovered, to be purposefully created by conforming to valued norms than to be passively observed as already existing.

Nor is this philosophy different in kind from "law" in the physical sciences. The neat mathematical law of gravity, for example, has never been observed in its perfection for objects on the surface of our earth. It is complicated by air resistance, and only as this is experimentally diminished, is the mathematical law approached. Laws are abstract statements of potential relationships which occur only under specified conditions and man, by modifying those conditions, makes the law operate, i.e., observes it in its relatively pure and exact form and not simply as a "tendency towards." Similarly, normality of incomes is a potential, mathematically-stated social law which may or may not be made operative in its pure form, depending on whether or not man modifies the prerequisite conditions. The tendency to normality of incomes is comparable to the tendency for objects to attract each other as their masses and inversely as the square of the distance between them. Each is a tendency imperfectly realized whenever other obscuring variables are present in the situation. The excellence of the law is proportional to (a) some measure of the strength of the tendency, i.e., the relative importance of the obscuring variables, and (b) our ability to control, or at least measure, each of these obscuring variables. This last test (b) is what makes the physical laws, in general, more excellent than social laws. The excellence of scientific "laws" is thus seen more as a function of the completeness of our knowledge of the phenomena than as a function of the nature of the phenomena themselves.

To avoid confusion, law as a "tendency," a summary statement of an existing situation with all its attendant irrelevant or complicating variables, and law as a principle, a statement of a potential situation abstracted out of its complicating variables, should be distinguished.

18. The population of the world for 1930 is estimated in the Encyclopedia of the Social Sciences (Ref. 38) at 1,901,380,000. Estimates for 1938 run up to 2.2 billions. The chances of exceeding a range of  $12.5\sigma$  in a normal curve are by Conrad and Krause's table (Ref. 37), .8 in 2 billion or .88 in 2.2 billion (i.e.,  $4 \times 10^{-10}$ ). This is less than one person, and dividing it by 2, to get the probability of a person occurring at either extreme of the curve, gives .4 persons in 2 billion (.44 in 2.2 billion) which is less than half a person.

19. A beginning of this research on the dispersions of human characteristics

has been made by Wechsler (Ref. 79) in collecting published distributions of physical, physiological, and psychological measurements of adults. His striking finding is that the ratio of the highest ability to the lowest ability averages 2 to 1, with a very small scatter from the lowest ratio of 1.45 : 1 up to the highest ratio of 3.23 : 1. His hypothesis is that this ratio of 2 : 1 represents a "natural constant." The units varied over pounds, inches, actions per minute, cubic centimeters, months of mental age, or vibrations per second. The human capacities measured were:

Weight of body	Blood pressure	S. reaction time
Weight of brain	Pulse rate	Audible pitch
Weight of heart	Vital capacity	Snellen acuity
Cranial capacity	Respiration rate	Intelligence quotients
Strength of grip		Binet mental age (both sexes for age 9)
		Memory span

He calculates the ratio of highest to lowest ability using the natural zero point of these scales as origins, and, by either dividing the highest score by the lowest (where the author reported these), or by dividing the mean  $+3\sigma$  by the mean less  $3\sigma$ . This latter procedure assumes (what is the crucial point to be tested in our hypothesis) that the range is  $6\sigma$ .

He supplements these data with a table of occupational piece work ratios between output of most efficient, and output of least efficient workers. Among the jobs were hair-trimming, loom operation, shoe bottom scoring, and polishing spoons. Here again the median ratio of the best to the worst operators was 2 : 1.

The twenty-two studies, which Wechsler summarizes, all had scales with natural zero points and populations under 4100 persons. Extreme deviates were often not reported. This indicates the directions in which further research might proceed to explore further the evidence which Wechsler has unearthed, for a "natural" range of human capacities.

20. This is equivalent to Von Wiese's formula for group behavior, stating that a societal process (P) is a function of attitudes (A) and the external situation (S):

$$P = A \times S$$

21. It should be noted that E, the tension ratio, is but a special case of S, a quantified societal situation.

In  $S = (T;I;L;P)$  the indicant in the tension theory is the ratio of the average desire to the total quantity of the value,

$$I = D/V \quad (\text{Eq. 37, Ch. V})$$

When the tension theory expresses a relation existing at any moment of time the exponent of T is zero. As geographical location is immaterial here, the exponent of L is also zero. Substituting  $T^0$ ,  $L^0$  and Eq. 37, Ch. V, into the S-equation, and noting that the generalized operational relation, symbolized by the semicolons, is here a multiplicative product,

$$S = PD/V = E \quad (\text{Eq. 34c, Ch. V})$$

which is the tension theory as defined by Eq. 34a, Ch. V. The S-situation is here a single index,  $(I)_E$ , which is simply denoted by the letter E, the equilibration or tension index.

22. Other indices of intensity of desire might be "money spent," an attitude test "score," a "discrepancy" between a prediction score (based on multiple regression weighting of previous achievement in high school, College Board examinations, aptitude tests, etc.) and marks in college, as a measure of motivation, or effort, relative to each student's ability, etc.

23. This "worth" should not be confused with other points of view, such as the worth of a college degree in the business world. Here V is a different one with its different P and D and consequent E.

The essence of analysis of a situation by means of the intellectual tool (Eq. 34, Ch. V) is to isolate and evaluate the factors relative to *one desideratum* at a time. This may be called the *field* of that particular desideratum. Difficulty in applying Eq. 34, Ch. V to various life situations is usually due to dealing with several different fields simultaneously, some of which may be vague and unexpressed, their corresponding P and D quantities consequently not clearly isolated in thinking.

24. This probability of passing is  $\frac{1}{P}$ , the ratio of those graduating to those entering college.

25. This ratio, E, may be modified by the units in which D and V are expressed, of course. It may require a factor within E to convert one unit into another as the constant .6745 converts sigma units into P.E. units. It will be comparable with another E only when D and V are converted into comparable units.

26. Throughout the statement of this tension theory of societal action, one is hampered for lack of concepts free of particular connotations limiting it to one specialized social discipline. Thus, "wants" instead of "desires" has an economic connotation, "values" is general in Philosophy but has ambiguity in distinguishing between "a value" meaning "an object desired," "a value" meaning in Economics the ratio of demand to supply, "a value" in Mathematics meaning a particular amount of some variable as it varies, etc. But any other term like "satisfactions," "goods and services," "things," "psychic objects," etc., would be still less precise and adequate. There is a need for a name for E more definite than "coefficient of value," "worth," "desire ratio," "tension ratio," "equilibration index," yet not involving money as "price" does. "Tension" connoting an attitude, a readiness to act, is behavioristically a necessary precondition, a cause of societal actions, and, thus interpreted, is to us the more useful concept for "naming" the "E" defined by Eq. 34a, Ch. V.

27. Recall the definition of D in Eq. 35, Ch. V and substitute it in Eq. 47b, Ch. V.

28. This conclusion may be re-derived as follows: E, as shown in Eq. 34a, Ch. V may be thought of as the psychological "price" for each unit of the value, for each friendship. Since  $P/V = 1$  by Eq. 47a, Ch. V for the case of unlimited

values, substituting this into Eq. 34a, and using Eq. 35, Ch. V gives:

$$E = \Sigma I_D / P (= {}^M I_D = D) \quad (\text{Eq. 47c, Ch. V})$$

an alternative form of Eq. 47b, Ch. V. This may be thought of as saying that the psychological "price,"  $E$ , of each friendship is the average desire of the people for friendships.

**29.** Refinements of this description, by considering interrelations between the quantities in the cells, may be ignored for the present, until it becomes clear whether the general theory will be useful to social scientists, or prove to be merely an ingenious but sterile manipulation of concepts and symbols.

## Chapter VI

### CORRELATION, I<sup>2</sup>

#### I. THE MEANING OF CORRELATION

##### A. Correlation in Quantic Terms

In our systematic exposition of the quantic classification of societal phenomena, visualized in the quantic solid, S. 33, Ch. II, the array containing indicators to the second power, I<sup>2</sup>, has now been reached.<sup>1</sup> This is the class of correlated characteristics.

Hitherto, one or more indicators which were merely aggregated or compared were considered (as shown by the matrical formula I<sub>i</sub><sup>+</sup>). We now proceed to a more penetrating level of observation achieved by the technic of cross-classifying two (or more) indicators in order to discover their causal relationships or other inherent connections more fully.

##### 1. A PRODUCT MOMENT OF TWO INDICES, ΣIJ/P

Thus for example, let I and J (and the class scripts, |<sub>I</sub> and |<sub>J</sub>) represent two indicants, each with a number of class-intervals, i<sub>I</sub>, j<sub>J</sub>. Taking a population, P, on which the two indicants have been observed, each class-interval of I (occupying a column) is subdivided into each class-interval of J (occupying a row). This cross-classifying can be equally well considered the other way around as subclassifying each class-interval of J into all the class-intervals of I. The essence of cross-classification is that it is mutual, that both indicants are independently observed. This fact is denoted by doubling the colon to symbolize that for each class-interval of one indicant there corresponds all the class-intervals of the other, *and vice versa*. The number of parties observed as belonging in any class-interval of the I characteristic and of the J characteristic is entered in the corresponding cell of the cross-classification table. This is the familiar correlation scattergram, or "scatteration" as it is sometimes called. (See

S. 7, Ch. II; S. 25, 26, 29, 33, Ch. VI; and S. 3, 5, Ch. XII, for examples.) Its descriptive formula is:

${}_iI :: {}_jI : P$  the correlation scattergram of two indicants among  
P persons (Eq. 1, Ch. VI)

This formula asserts that for every one of the  ${}_i|$  class-intervals cross-classified with all the  ${}_j|$  class-intervals there is a corresponding frequency of persons. This is the simplest and perhaps the most common form of societal correlation. For complete generality it should be observed that the units of frequency which are persons in Eq. 1 may be plurels, or dates, or areas, or even other indicators, i.e., that the cross-classified entities may be indices.<sup>2\*</sup> Since through all these forms, the cross-classifying of two indices,  $(I) :: (I)$ , is the common core of their descriptive formulae and as this is denoted by the second power in the quantic formula,  $I^2$ , that is taken as defining the category of societal relationships which bear the concept-name of "correlation."

The question may be asked why cross-classification of indicants involves a quantic of  $I^2$ , whereas aggregation of them involves a quantic of the first power only,  $I^{+1}$ . The answer is that, while the scatteration is an aggregation of numbers, the summary of these which is the correlation coefficient is a multiplicative *product* of the two indicants. Thus, in rectilinear correlation the Pearson product-moment formula,

$$r = \frac{\Sigma IJ}{P\sigma_I\sigma_J} = \frac{\Sigma I_\sigma J_\sigma}{P} \text{ coefficient of correlation} \quad (\text{Eq. 3, Ch. VI})^3 \dagger$$

is the average of the I times J products (each expressed in sigma units). The quantic formula for the correlation coefficient (Eq. 3, Ch. VI) is  $I^2P^{-1}$ , which is thus distinguished by the sign of the exponent on P from the quantic of the scattergram of Eq. 1, which is  $I^2P^{+1}$ .

A statistical formula is an operational definition of a concept. It tells the computer what to do to get the entity defined. The fact that the accurate formulae<sup>4</sup> for calculating correlation involve a product of indicants means that an indicant to the second power is an essential part of the definition of correlation.

\* For Eq. 2, Ch. VI, see notes at end of the chapter.

† For Eq. 4, Ch. VI, see notes at end of the chapter.

## 2. A SCALAR PRODUCT OF VECTORS, $\bar{I} \bullet \bar{J}$

Vectorial algebra corroborates the quantic formula of  $I^2$  for correlation. The scalar product of two vectors is defined by mathematicians as the product of the lengths (scalar aspects) of the two vectors times the cosine of the angle between them. If each vector is taken one unit in length their scalar product simplifies to "the scalar product of unit vectors," which is the cosine of the angle and is also the correlation coefficient between the two variables represented by the vectors. The correlation coefficient is a scalar product when sigmas are the units. The scalar product (or dot product as some mathematical textbooks call it) is denoted by the heavy dot,

$${}_s I \bullet {}_s J = r_{IJ} \quad (\text{Eq. 5a, Ch. VI})$$

This relation, Eq. 5, states the identity of a scalar product in sigma units in S-notation with a correlation coefficient in conventional statistical notation.<sup>5 \*</sup>

As shown in diagram S. 13, Ch. VI the scalar product of two vectors is geometrically the area of the parallelogram defined by one vector and the normal to the other. At the upper limit when the parallelogram becomes a rectangle the cosine, or  $r$ , becomes unity and the scalar product of the two collinear vectors becomes identical with the ordinary arithmetic product of two numbers represented geometrically by a rectangular area. Thus the scalar product is the more general case including the ordinary arithmetic product as a special case. As most phenomena in the social sciences have a degree of correlation to each other of less than unity, the scalar product is the more useful and generally applicable description of their relationships than the arithmetic product. Consequently to define the second power of the indicator in S-theory the scalar product has been chosen. This scalar product is statistically computed via the correlation coefficient.

### *B. Correlation in Terms of Probability and Causation*

While correlation is operationally defined by the scalar product of two indices, yet a deeper insight into its meaning involves its relation to the concepts of "probability" and "causation." For the sociologist in common with all scientists, as stated in the

\* For Eqs. 6a-b and 5b-c, Ch. VI, see notes at end of the chapter.

introductory chapter, wants to understand, predict, and control the phenomena in his field. For such understanding and prediction, knowledge of probabilities, correlations, and causes are essential. Our hypothesis on this point is that "probability," "correlation," and "causation" are compounds, in relationships definable by formulae below, of our more elementary concepts of indicators, time, and population.

## 1. SIMPLE PROBABILITY AS A RATIO OF FREQUENCIES OF INDICES

A simple probability is a ratio of frequencies of an all-or-none indicant ( $I^{0-\infty} = {}^{10}I$ ). The units of frequency may be dates, persons, or some other index. The probability index is defined as the ratio <sup>6</sup> of cases of some quality to cases <sup>7\*</sup> -having-plus-cases-not-having that quality. If cases of that quality are symbolized by  ${}^{21}I^0$  and cases without it as  ${}^{21}I^{-\infty}$ ,

$${}^{21}I^0 / ({}^{21}I^0 + {}^{21}I^{-\infty}) = i / (i + j) = \text{a probability} \quad (\text{Eq. 7a, Ch. VI})^8$$

## 2. NORMAL PROBABILITY AS A SUM OF SIMPLE PROBABILITIES

The next step in compounding occurs when an observed fact is the sum of more than one all-or-none attribute. A fact such as a person's measured height or intelligence score may be considered as due to a large number of separate, qualitative, determining elements, which on the average are more or less equivalent to each other.

The probability of such a sum is a coefficient of the binomial expansion which approaches the normal distribution curve as a limit.<sup>9</sup> † Reversing the reasoning <sup>10</sup> a normally distributed characteristic may be thought of as a resultant of  $n$  hypothetical determining elements which are the cases of the attribute  ${}^1I^0$ . (These "elements" will be discussed more fully below.)<sup>11</sup> Once we observe that the indicant measuring a characteristic is normally <sup>12</sup> distributed and find its mean and sigma, prediction becomes possible with the aid of the normal probability table, as discussed in Chapter V. Thus it is evident that attribute ratios measure simple probabilities and sums of these produce distribution curves, which enable one form of prediction in the social sciences.

\* For Eqs. 7b-c, 8a-b, and 9, Ch. VI, see notes at end of the chapter.

† For Eqs. 10-10b, Ch. VI, see notes at end of chapter.

## 3. CORRELATION AS A PRODUCT OF PROBABILITIES

The next step in compounding probabilities is the product of two or more probabilities—and this is correlation. In its simplest form it is the product of two different all-or-none indicants.<sup>13</sup>

This last statement, often called the law of joint probability, is that the probability of a simultaneous or successive occurrence of two or more events is the product of their separate probabilities. If the probability of one characteristic is  $p$  and of a second  $p'$  (and  $q$  and  $q'$  are the complementary probabilities of their not occurring, i.e.,  $p + q = 1 = p' + q'$ ), then the probability of both characteristics occurring together is  $pp'$ . This is more fully shown in the diagram:

## JOINT PROBABILITIES OF TWO TWO-CATEGORY CHARACTERISTICS

Characteristic, I				
	Present	Absent $qp'$	Present $pp'$	Sum $p'$
Characteristic I'	Absent	$qq'$	$pq'$	$q'$
	Sum	$q$	$p$	$1 = p + q$ $= p' + q'$

These expected, or theoretical relative, frequencies in the cells represent a zero degree of correlation between the two characteristics. But in proportion as the actually observed frequencies deviate from the theoretical ones correlation increases, until it reaches its maximum at  $r = 1.00$  when all the frequencies are in one set of diagonal cells and the other cells are empty, i.e.,  $pp' + qq' = 1$ ,  $qp' = 0 = pq'$  (for positive correlation).

This fourfold correlation is not just an item in the statistician's bag of stunts. It is implicitly involved in almost every judgment a sociologist makes. If he asserts, "Adolescents are emotionally unstable," he asserts a correlation between the two all-or-none indicants, "adolescent, non-adolescent" and "unstable, stable," affirming that the "adolescent and unstable" joint category has a larger proportion of cases than expected by the law of joint probability. The degree to which any such generalization holds true in any defined population can be formally determined by calculating the appropriate form of correlation or contingency

coefficient, which summarizes the relationship revealed in the diagram above. The difference between sociologists who use and who do not use statistical technics is not in their dependence on quantitative reasoning but rather in the explicitness of such reasoning. The latter group assert such generalizations as "adolescents are unstable" without determining the degree of the correlation or the reliability of determining the classes "adolescents" and "unstable." While the statistical group realize that, for the generalization to be a scientific one, they must find some precise measure of the extent to which characteristics such as "adolescence" and "instability" vary together and of the reliability of observing those characteristics.

The logic of the diagram above extends beyond all-or-none indicants to ordinal and to cardinal indicants. Here the correlation merely becomes rank  $r$  for ordinals and then the still more exact Pearson  $r$ , or ratio  $\eta$ , for cardinals in harmony with our theory of increasing precision of observation presented in Chapter III in discussing the class-interval script of indicators. For cardinal indicants the diagram above becomes the usual correlation scattergram (illustrated in S. 1, 2, 5, 20, 25, 26, 29, 33, Ch. VI). In another direction, if the characteristics are qualitative ones represented by attributes with subclasses the diagram above merely becomes the contingency table shown in S. 11 and 27, Ch. VI. The contingency coefficient summarizes the degree of deviation of the observed frequencies from the theoretic frequencies that are the most probable ones by chance. This contingency coefficient becomes identical with the correlation coefficient when the two characteristics are normally distributed and expressed in many class-intervals. In this case the theoretic probability in any cell, i.e., for any particular magnitude of each indicant, is given by the height of the column in that cell up to the normal correlation surface diagramed in S. 5, Ch. VI.

The conclusion is then that correlation is a product of at least two probabilities.<sup>14</sup> The many forms of correlation vary with the precision of the indices that are correlated and with their form of distribution. But all forms of correlation are a function of at least two indices and their relative frequencies, as stated in Eq. 2, Ch. VI. Thus the thesis is borne out that correlation, like probability, is a compound of the more elementary concepts of indi-

cators and population in particular relationships which are specified by their formulae.

#### 4. CAUSATION AS CORRELATION PLUS TIME SEQUENCE

There remains to be exposed the relation of the concept of "causation" to "probability," "correlation," and to the basic indices of S-theory. Causation may be analyzed into two essential ingredients: antecedence in time and a probability of sequence. For A to be a cause of B, A must precede B in time, and there must be a probability from past experience<sup>15</sup> that B will follow A. If this probability is perfect, A may be the sole cause of B, if it is intermediate between zero and unity, A may be proportionately a partial cause of B. Causation is a subcategory of correlation, being a correlation where one variate precedes the other in time. (Of course, both may be continuous processes; any antecedence will be evident whenever the correlation rises when corrected for some lag.)

Complete causation as Mill and Pearson long ago pointed out is uniform antecedence. But while time sequence and a probability of it are necessary tests of causation they are not sufficient tests, for a phenomenon which is not a direct cause may pass both tests. Thus day precedes night with perfect probability in human experience, yet the day does not cause the night, but rather both are effects of a causal background, namely, an opaque spherical earth rotating in a field of uni-directional light rays coming mostly from one source. Another instance is the observation of a correlation of .84 between income and score on a hygiene scale in a village and city population in Syria. Case study evidence shows some probability, i.e., some cases, where low income prevented medical care and high income enabled cleanliness, etc. But similar evidence was found showing that poor hygiene and consequent ill-health decreased earnings. Hygiene and income, then, interacted, each being at least a partial cause of the other. But more fundamentally, the whole rural-urban background explained the correlation, for in such a wide range almost any two characteristics showing low correlation in rural groups alone or in city groups alone showed high correlation in a combined rural and urban sample. The education, contacts, medical facilities, and the rest of the differences between the rural and urban population seem

to have produced both the income disparity and the hygienic disparity more than either directly caused the other.

Such considerations lead to the conclusion that, whenever probable sequence of A and B is both a necessary and sufficient test, A and B are directly causally related, but whenever the probable sequence of A and B is a necessary but insufficient test, A and B may be indirectly related through some third causal condition in common. The sufficiency is discoverable by the human ability or inability to duplicate or reinstate A and observe that B tends to follow.<sup>16</sup> Only when daylight and night can be duplicated on a rotating sphere in a field of light is the knowledge of the essential conditions of causation complete; for then the fully sufficient as well as the necessary causal conditions are known.

Only when a phenomenon, such as hygienic status, can be varied at will under different attendant conditions is its causation fully understood. This test of duplicating and varying the conditions and finding the effect to follow selects the true causes out of all the frequently concurrent conditions. Antecedence, probability, and the technic of experimental duplicating under varying conditions, then, would seem the necessary and sufficient operational definition<sup>17</sup> of causation. Whenever we cannot experimentally duplicate or statistically isolate, we can only assert a causal connection between correlated variables<sup>18</sup> which may be indirect through other factors in common.<sup>19</sup>

Further discussion of causation, including a formula for it, will be deferred to Chapter XI on Societal Forces in discussing the time sector.

### *C. Correlation in Sociological Terms*

#### 1. CATEGORIES DEFINED BY THE QUANTIC FORMULA

The sociological reader will probably take exception to making the statistical and methodological concept of "correlation" a fundamental category in a systematics of Sociology, on a par with "interaction," "societal processes," "forces," etc. This is done because clear-cut consistency with the theory on which this system is based requires it.

The systematics in this volume is based on the quantic formula:

$$S^s = T^t; I^i; L^l; P^p \quad (\text{Eq. 11, Ch. VI})$$

The combinations of the four types of indices and their exponents constitute the quantic classification (see Ch. II) and define the chief sociological categories and corresponding chapters devoted to their exposition. In this classification there are combinations such as,  $T^{-1}IP^2$ , which is found to define the accepted sociological concept of "interaction"; there are  $T^{-1}I$  and  $T^{-1}P$ , which define "societal processes"; there are  $T^{-2}IP$ , which is found to define a "societal force," and  $T^{-2}IP^2$ , which defines "societal control";  $L^{-2}P$ , which defines population "density";  $I^0P$ , which defines a "plurel." There is also  $I^2P$ , which defines "correlation." Adherence to the quantic hypothesis requires that correlation then be treated as co-ordinate with the other concepts defined by the quantic formula.

But aside from the quantic formula, why are not the relations of characteristics of people, the patterns of their culture, as legitimate a sociological category of content as the relations of people to each other in respect to some characteristic? Or as legitimate as the temporal and spatial relations of changing characteristics, as in concepts of "lag," "culture diffusion," "gradients," etc.? The historical accident that "correlation" of characteristics of people was reduced to formulae by statisticians long before "interrelations" of people, which is just beginning to be reduced to formulae, is surely insufficient logical ground to deny the former and accept the latter as denoting sociological content.

No attempt will be made here to list or classify the large mass of correlations of societal phenomena reported in the sociological literature. The attempt would require going into far greater detail in discussing the many intercorrelations between all pairs of sociological fields or topics than would be included in a volume of systematics. The major part of every volume on specialized subfields, such as the Family, is devoted to tracing such correlations among its characteristics and changes. Instead of a catalogue of correlations by content correlated, general aids to interpreting correlation will be offered which may find application in studies of diverse content. (See appended graphs S.1 through 24, Ch. VI.) This is in line with the policy of limiting this volume,

as stated in the introductory chapter, to the definition of general Sociology as the study of the "general characteristics common to all classes of social phenomena."

## 2. METHODOLOGICAL CONCEPTS IN A SYSTEMATIC SOCIOLOGY

These "general characteristics" are necessarily partly methodological and statistical, but only partly so. Operational definitions of sociological concepts necessarily integrate content and methodology. Statistical concepts are operationally defined by their computational formulae, and the superiority for science of using operational definitions is one of our basic assumptions as stated in Chapter I. It is the avowed aim of this volume to integrate sociological theory and statistical theory, to promote the wedding of Mathematics and Sociology. Statistics has outgrown a collection of technics for marshaling sociological facts. Statistical theory can usefully give precision to sociological theory as in defining "processes," "forces," etc. Sociological theory should be organized around more objectively determinable concepts, as for example, in our defining of social "problems" in terms of negative standard deviations of some desideratum. This volume is inadequate as a statistical text. It assumes the reader knows some statistical theory and is ready to apply it. On the other hand, this volume is also inadequate as a complete systematics of Sociology. It is expressly limited to the quantified aspects of Sociology. Within this field it seeks to offer an orderly framework towards building a more objective and exact science of Sociology, and within this field the concept of correlation is an essential piece of that framework.

### *D. Correlation in Geometric Terms*

#### 1. THE CORRELATION COEFFICIENT AS THE COSINE OF THE ANGLE BETWEEN TWO VECTORS

In discussing the geometric interpretation of S-theory it has been stated that quantities are represented geometrically by lengths of lines and qualities are represented geometrically by angles or directions. Every index is representable as a vector, which by definition is a directed line. The chapters on Indicators and on Plurels discussed units for determining the length of their

vectors and discussed classifications for determining the number of their qualitatively different vectors. We now come to examine the degree of qualitative difference, the size of the angle between any two vectors. This is measured by the correlation coefficient which may be converted from a table of cosines into degrees of angle. In proportion as the vectors of two indices make an angle varying from  $0^\circ$  to  $90^\circ$  the two indices vary from identity through decreasing similarity to the limit of complete qualitative difference or independence of each other.

For an example, what is the degree of qualitative similarity between College Entrance examinations and Freshman year standing? The question is answered in S. 34, Ch. II,<sup>20</sup> where in one large university the observed correlation of .40 means an angle of  $66.4^\circ$  between their two vectors. This graph compares the similarity of other indicants such as individual examinations and high school record by comparing the sizes of their angles. The practical purpose in this graph is to find those indicants which are most like the criterion indicant to be predicted, in order to predict it more accurately. From this graph it is evident that the multiple correlation of a team of predictor indicants is more similar to (i.e., correlates higher with and therefore predicts better) the criterion of later academic achievement in college than any one of the predictor indicants alone. In causal terms, since there is correlation between an antecedent achievement of students and a subsequent achievement, it can be asserted that there is a causal connection which, however, in this case is probably indirect. Until fuller experimental duplication of the causes producing college achievement is made, it can only be asserted that the earlier indicants and the later indicants measure some causal factors in common, such as perhaps the academic ability, industry, and academic opportunity of the students.

## 2. INTERCORRELATIONS AS A SYSTEM OF DIRECTION COSINES IN $n$ -SPACE

These principles can be extended from pairs of indices to situations involving  $n$  indices, whether of the same or of different sectors. Whenever it is possible to secure all the intercorrelation coefficients between all possible pairs of indices, these coefficients interpreted as direction cosines specify all the angles between all

$n$  vectors representing the  $n$  indices. Usually these angles cannot be all consistent in a three-dimensional space. All the angles are always consistent, however, in a space of  $n$  dimensions. This situation is diagramed in S. 35, Ch. II which shows a sheaf, or ray, of vectors springing out from a common zero origin. The  $n$  vectors (here  $n = 8$ ) define  $n$  dimensions of societal space and the temporal, indicatory, spatial, and populational vectors define those four sectors of that space in this situation. Note that the angles between vectors in any sector as well as between vectors in different sectors may be of any size from  $0^\circ$  up to  $90^\circ$ . Whenever the angle between vectors in different sectors is zero those vectors are collinear, i.e., become one, and one which lies on the boundary between the sectors. Thus, the correlation between the size of populations,  $P_p$ , and some characteristic,  $I$ , given in per capita units, such as food consumed stated in colloquial units of "mouths to feed," is obviously unity.

$P \bullet I = 1$ . These two vectors coincide at the boundary of the populational and indicatory sectors. Again the monthly frequencies of temporal events, such as a doctor's calls, might correlate perfectly with the monthly number of patients seen if he sees one and only one patient in a call. The vector representing the events, "calls,"  $T$ , coincides with the vector representing the population, "patients,"  $P$ , and marks one boundary line between the temporal and the populational sectors ( $T \bullet P = 1.0$ ). But such instances are unusual. Usually vectors from different sectors show larger angles. Angles of  $90^\circ$  are very common, as in many data the amount of the characteristic, the size of the population, the date, and the area are not intercorrelated.

In this geometric representation of societal  $n$ -space, persons (or plurels) may be accurately locatable and their distances from each other can be calculated. Mr. X is a Republican, Presbyterian, from Montclair, New Jersey, in the \$10,000-a-year-income class, married, childless, etc. Each of these six characteristics is a vector, and with fuller data to determine the correlations (or contingencies in qualitative data) the angles of the sheaf of vectors become known and Mr. X as a point on each vector has a definite location.

Previous sociologists have used the concept, as Sorokin did in *Social Mobility* (Ref. 68), but only in an analogical way without

developing the mathematical technics for actually determining distances and directions in this space. The detailed steps in computing distances in sigma units and angles from correlations are more fully developed in textbooks of statistics and of matrix algebra. The function of a systematic Sociology such as the present volume aims to be is to point out where all classes of sociological phenomena fall, leaving their detailed development to more specialized volumes.

## II. THE ANALYSIS OF CORRELATIONS

### *A. Transformation of Indices*

For the beginner in dimensional analysis, orthogonal dimensions may be thought of as mutually perpendicular co-ordinates (called Cartesian co-ordinates). In the three dimensions of physical space, these can be visualized as the three edges of a rectangular box, or of a room, which meet in one corner. Visualizing breaks down beyond three dimensions, although conceptually it can be extended. A helpful transition to conceiving of more dimensions is, to some students, to think of time as a fourth dimension. The three edges of the box exist in an invisible time dimension. They move along a time co-ordinate from the past through the ever sliding present on towards the future. Mathematically by means of matrices and direction cosines (see S. 15 and 24, Ch. VI) any number of dimensions can be handled as precisely as three dimensions, though with more computational labor. A table of intercorrelation coefficients of  $n$  characteristics, each correlated with each of the others, provides such a matrix of direction cosines specifying all the angles between the  $n$  vectors in the  $n$ -space which they define.

It is often desirable to re-express a vector in terms of two standard co-ordinates of the plane in which the vector lies. This can be done by writing the equation of a straight line for that vector in terms of the two co-ordinates  $x$  and  $y$

$ax + by = 0$  the equation of a straight line (through the origin)  
(Eq. 12, Ch. VI)

This "transforms" the vector into a certain amount of the  $x$  co-ordinate and a certain amount of the  $y$  co-ordinate. This

transformation can also be described as breaking up the resultant vector into an x "component" and a y "component." Since these x and y components are uncorrelated (by definition they form an angle of  $90^\circ$  whose cosine is zero), they fulfill the canon of classification which requires that the categories into which phenomena are resolved shall not overlap. These components are the mathematicians' and physicists' "components" of resultant vectors.

In Psychology these components have been called "factors" by Spearman, Thomson, Thurstone, and others, or "components" more recently by Hotelling, Kelley, and others. "Factors" imply multiplicative parts of a whole product, whereas these components are added together. Consequently the term "component" will be used here to denote the variables into which a given set of intercorrelated variables are resolved. Geometrically viewed the components are Cartesian co-ordinates,<sup>21</sup> used as a standard frame of reference for describing the observed vectors. They may also be viewed as percentages of elements common to several indices. (See S. 7 and 8, Ch. VI.)

In general, it is always possible to re-express or transform  $n$  vectors into  $n$  co-ordinates by means of multiplication of matrices of correlation coefficients from equations such as Eq. 12, Ch. VI generalized.<sup>22</sup> This is a technical and rapidly growing field offering much promise of future importance.<sup>23</sup>

The utility of such transformations is twofold: first, it re-expresses the  $n$  observed and overlapping characteristics in standard non-overlapping characteristics; and second, it is often possible to re-express the observed characteristics in *fewer* non-overlapping characteristics, thus achieving the parsimony aim of science.

### *B. Reduction of Indices to Components*

#### 1. DECREASE OF THE NUMBER OF DIMENSIONS BY RANK OF THE MATRIX

Under certain conditions, it is possible to transform  $n$  vectors into fewer co-ordinates,  $n'$  in number, that is, to re-express  $n$  components in a smaller number,  $n'$ , of subcomponents,  $n > n'$ . The matrix of intercorrelation coefficients which is of order  $n \times n$  is then said to be of "rank"  $n'$ . This means that all the  $n$  vectors

lie within a space of  $n'$  dimensions. The critical condition for thus reducing the number of dimensions is that all orders greater than  $n'$  of the matrix of intercorrelations when treated as an algebraic determinant must equal zero. (See S. 21, Ch. VI.) In practice these determinants do not always have to be evaluated. The components may be computed successively in order of importance until a point of diminishing returns is reached, when further components are about the size of the sampling errors and so are not worth calculating. For practical purposes the number of components computed up to this point is the rank of the matrix and the sufficient number of dimensions within the limits of error of the data.

## 2. THE REDUCTION TO PRINCIPAL AXES

To date several alternative methods of analysis have been developed for transforming the observed indicants into subcomponents.

One method is that of principal axes for which Hotelling (Ref. 29) developed an iteration technic and Kelley (Ref. 34), a simpler technic of rotating the sheaf of observed vectors in one plane at a time to compute axis after axis. The principal axis can be visualized in S. 2, Ch. VI as the major axis of the ellipse representing the scatteration.

The second principal axis is the next longest axis at right angles to the first—the minor axis of the ellipse in the two-dimensional scatter of S. 3, Ch. VI. For a three-dimensional illustration see S. 41, Ch. VI, which shows the “swarm” of points representing the intercorrelation of three predictive indicants of college achievement. In this ellipsoidal swarm the principal axis is the longest one possible, the next axis is the next longest, which is perpendicular to the first, the third is the longest residual axis which is perpendicular to the two previously extracted longer axes. For  $n$  dimensions the procedure goes on similarly. The components are orthogonal and are extracted in order of their contribution to the total variance of the “swarm.”

## 3. REDUCTION TO CENTROIDS

A second method is that of the center of gravity, or centroid, developed by Thurstone. (Ref. 77.) By this, the first component

is the centroid to a cluster of vectors, a sort of average of them. When the vectors are all unit vectors (i.e., reduced to sigma as the scalar unit) the sheaf can have their terminal points represented on the surface of a sphere. The centroid is the central point as graphed in S. 15, Ch. III and S. 21, Ch. VI. The second centroid is the average of the remaining vectors after eliminating the first, and so on for further centroids.

Thurstone reports (Ref. 76) a study in which 60 trait terms (adjectives) were intercorrelated with the aid of 1300 raters and reduced to five components or centroids of clusters of traits. In another study he reports that 45 psychotic traits were reducible to five subcomponents—a manic, a catatonic, a cognitive, a hallucinatory, and a depressive cluster.

These centroidal components need not be orthogonal. They must be rotated before they can be interpreted in terms of the observed data. The orthogonal principal axes type of component would seem to be the more readily interpreted. Much further research, however, is necessary for determining further properties and utilities of both types.<sup>24</sup>

A special case of considerable importance is that of one general component and  $n$  specific ones in  $n$  observed components. This is historically the earliest and for many years the chief analysis into what its discoverer, Spearman, called the factors  $g$  and  $s$ . (Ref. 69.) Whenever the intercorrelations are in equiproportion, a term defined and tested by the vanishing of all tetrads of them (i.e., second-order determinants) within probable error limits, the  $n$  indicants can be said to be composed of one component general to all the observed indicants and  $n$  components each specific to one indicant. This condition is alternatively stated as that the matrix of intercorrelations has a rank of one. This condition is frequently found among homogeneous sorts of complex indicants such as verbal tests of intelligence.

#### 4. ANALYSIS INTO COMPULSORY COMPONENTS

The methods of analysis above are optional in the sense that alternative analyses into other patterns of components are possible and the investigator chooses, within certain limits, the principal axes pattern, or centroid pattern, on grounds of parsimony, simplicity of pattern, or ease of interpretation.

But a further compulsory type of analysis has been developed by Thomson and Thompson (Refs. 72 and 74) and others, as shown in S. 43. This can be neatly summarized in a determinant of the intercorrelations, which, when the successive integers are used in the diagonal cells, according as it is positive or negative upon evaluation, tells whether components ("factors" in S. 43) above or below a specified order are necessary. "Necessary" here means that it is physically and mathematically impossible to secure such correlation coefficients without the necessary orders of components. The "order" of a subcomponent is the number of indicants of which the component is a part. This type of analysis yields orthogonal components whose number, however, will usually exceed the number of observed indicants, since at each order  $q$  (where  $q$  takes all values in turn from 1 to  $n$ ) the number of components is  ${}_nC_q$ , the combinations of  $n$  things taken  $q$  at a time. These are the binomial coefficients. These are also the number of unknown components. But the number of known relations from which to solve is limited by the number of correlation coefficients,  ${}_nC_2$ , and is actually less as these are not all independent. Since to solve for any number of unknowns the same number of known and independent relations must be available, this type of analysis is limited in practice to certain orders only. A further limitation seems to be that the magnitudes of observed intercorrelations are seldom such as to show, by this determinantal test, a compulsory general component. At least very little application of this method to societal data seems to have been made to date.<sup>25</sup>

## 5. REDUCTION OF PLURELS INTO "TYPICAL PERSONS"

Still another type of analysis, called Q-technic by Stephenson (Ref. 70), is to calculate components in the populational sector of societal space instead of in the indicatory sector as above. As suggested in S. 24, Ch. VI, a correlation is calculated between the  $n$  indicants of two persons (instead of between the two indicants of  $N$  persons, as in the "r-technic" of the previous methods). All possible pairs of persons yield a matrix of intercorrelations,  $N^2 - N$  in number. Either the principal axes or centroidal methods may be applied to this matrix, yielding *types* of persons as components. Thus, the first principal axis would be inter-

preted as an hypothetical most typical sort of person; the second axis would be the next most typical person whose characteristics were independent of those of the first most typical person; and so on. Much research remains to be done in elucidating and trying out this type of analysis into populational components.

### *C. Reduction of Components to Elements*

#### 1. CORRELATION AS A PERCENTAGE OF COMMON ELEMENTS

There is a possibility of pushing the analysis of components still further into simpler elements. As noted above, a normally distributed variable can be considered to be a function of a large number of small interchangeable elements, which act according to the law of the probability of sums in giving rise to the normal frequencies.

The simplest way of thinking about such elements as possibly constituting components and indices is in terms of the formulae for the correlation coefficient as a percentage of common elements. By this formula the correlation coefficient is the geometric mean of the two proportions that the number of elements common to two indices is to the number of elements in each index.

$$r_{IJ} = \sqrt{\frac{c}{c+i} \times \frac{c}{c+j}} = \text{correlation as a percentage of common elements (see S. 7, Ch. VI)}$$

(Eq. 13a, Ch. VI)

Here  $c$  is the number of elements common to the indices  $I$  and  $J$ , and  $i$  and  $j$  are the number of elements specific to each index.<sup>26 \*</sup>

#### 2. THE HYPOTHESIS OF EPSILON ELEMENTS

When the properties of the elements are known, or can reasonably be assumed, further consequences can be adduced. Suppose the elements are all-or-none attributes. Let us assign an arbitrary numerical value of 1 to the presence of the attribute and of -1 to its absence.<sup>27</sup> Then the mean of the equally-frequently present or absent characteristic is zero and its standard deviation (and hence its variance) is unity.<sup>28 †</sup>

Let such unit-elements be designated by the Greek letter epsilon,  $\epsilon$ . When a sample of these  $\epsilon$  elements are added together

\* For Eqs. 13b-c, Ch. VI, see notes at end of the chapter.

† For Eqs. 14-19c, Ch. VI, see notes at end of the chapter.

to form an index, the variance of the latter is given by the number of the elements in it. In two correlated indices their covariance is given by the number of elements common to both.

Since the covariance and the two variances ordinarily are observed quantities, Eqs. 18 and 19 give three knowns from which the three unknown numbers of elements,  $c$ ,  $i$ , and  $j$ , may be readily determined. This is true whenever the indices  $I$  and  $J$  are known or assumed to be composed of such  $\epsilon$  elements. But the difficulty is to know when such an assumption is justified. Because it is proved that elements give indices whose variances tell the number of elements, the reasoning cannot then be reversed (except to form an hypothesis) to hold the converse statement true that an observed variance is the number of  $\epsilon$  elements. Even if they were, the units in which observed components are measured may or may not be the same as the unit-elements,  $\epsilon$ . These statements may be readily verified by calling pennies  $\epsilon$  elements, with heads = +1, and tails = -1. Indices may be artificially synthesized by tossing groups of pennies, leaving  $c$  pennies to lie over for the common component from one toss which gives a case of the  $I$  index, to a second toss which gives its paired case of the  $J$  index. When variances and covariances from a large number of tosses have been secured, the number of pennies actually used for  $c$ ,  $i$ , and  $j$  will correspond (within standard error limits) to the theoretical  $c$ ,  $i$ , and  $j$  calculated from Eqs. 18 and 19. But suppose the observed variances were expressed in some transposed scale, such as multiplying by  $a$ . This would obscure the fact that the observed indices arose from elements. Alternatively we may say that elements whose sigmas are  $\epsilon\sqrt{a}$  would produce such indices. When the observer starts with given sociometric data, expressed in any convenient units, the problem is to determine the values of  $a$  (if they are not unity) for both  $I$  and  $J$ .

Difficulties such as these prevent the hypothesis of  $\epsilon$  elements from being universally and readily applied to sociometric data. It is possible that research will overcome these difficulties and enable the reduction of indices to elements which have the properties of  $\epsilon$ , or other properties, and open a fruitful field of sociometric analysis.

Indices of the  $T$ ,  $I$ ,  $L$ ,  $P$  kind in  $S$ -theory may be correlated and hence have overlapping or common elements. Orthogonal

components are by derivation and definition uncorrelated and hence have no elements in common. Thus one of the purposes of analyzing indices into orthogonal components is to eliminate overlapping elements and have each composed of a unique and different set of elements.

The analysis of societal phenomena by S-theory and this hypothesis then is a hierarchy of concepts as follows: any societal situation (i.e., convenient section of phenomena) is analyzable into its temporal, indicatory, spatial, and populational indices, many of which are intercorrelated in a complicated network of relations. These correlated indices can always be re-expressed in terms of uncorrelated components which, under certain conditions, may be fewer in number than the number of the observed indices. These components finally may be resolvable into numerous small, independent, homogeneous elements. Analysis such as this may prove to have as much importance for sociological data as analysis of physical compounds into molecules and atoms. At the present stage of research, however, this is a possibility but far from an actuality.

### 3. ANALYSIS OF THE GENERAL COMPONENT, $g$ , INTO ELEMENTS

In the case of equiproportional intercorrelations defined by a matrix of rank one, where the  $n$  indicants are analyzed by Spearman's technic into a general component,  $g$ , and  $s$  specific components, an alternative or further reduction to elements has been shown to be possible by Thomson, Garnett, Dodd, and others.

#### EQUIPROPORTIONAL INTERCORRELATIONS

a matrix of rank one  
(all tetrads vanishing within P.E. limits)

Tests					Array coefficients					
	a	b	c	—	n				—	
a		$r_{AB}$	$r_{AC}$	—	$r_{AN}$	$r_{AG}$	$r_{BG}$	$r_{CG}$	—	$r_{NG}$
						$(r_{AG}^2)$	$r_{AG}r_{BG}$	$r_{AG}r_{CG}$	—	$r_{AG}r_{NG}$
b	$r_{BA}$		$r_{BC}$	—	$r_{BN}$	$r_{BG}r_{AG}$	$(r_{BG}^2)$	$r_{BG}r_{CG}$	—	$r_{BG}r_{NG}$
c	$r_{CA}$	$r_{CB}$		—	$r_{CN}$	$r_{CG}r_{AG}$	$r_{CG}r_{BG}$	$(r_{CG}^2)$	—	$r_{CG}r_{NG}$
—	—	—	—	—	—	—	—	—	—	—
n	$r_{NA}$	$r_{NB}$	$r_{NC}$	—		$r_{NG}r_{AG}$	$r_{NG}r_{BG}$	$r_{NG}r_{CG}$	—	$(r_{NG}^2)$

=

(Eq. 20, Ch. VI)

(Summarized in Ref. 17.) Each of the equiproportional correlation coefficients is the product of two array coefficients which are the correlation coefficients between the general component and the component of that array. This may be clarified in the pattern on page 326.

Since  $g$  is an additive part of each index (which is a sum of  $g$  and  $s$  alone),  $r_{IG}^2$  is the proportion that  $g$  is of the index  $I$  (by Eq. 13b, Ch. VI), and  $r_{IJ}$  is  $r_{IG}r_{JG}$ , which is the geometric mean of two such proportions in accordance with Eq. 13a, Ch. VI. In terms of elements this means that the variance of each index, ( $I$ ), can be split into a proportion,  $r_{IG}^2$ , of general (i.e., completely common elements) and a complementary proportion,  $r_{IS}^2$ , of specific elements.<sup>29</sup> \*

While the proportions of these elements are determinable, their absolute number is not determinable, owing to the fact that there is one more unknown component ( $n + 1$ ) than the  $n$  known indices. Different indices are observed to have different amounts of  $g$ , i.e., differing proportions of the generalized elements. In verbal tests of intelligence, where equiproportion is most often observed, those tests which are "most highly saturated with  $g$ " are said to measure general intelligence best.

In terms of probability,  $r_{IG}^2$  is the probability that any element in test  $I$  will be general to the other tests. The probability of two tests having common elements is, by the law of joint probability, the product  $r_{AG}^2 r_{BG}^2$  of the independent probabilities. This is the index of determination,  $r_{AB}^2$ , whose square root is the observed correlation coefficient,  $r_{AB}$ . The equiproportional intercorrelations are thus the square roots of the theoretical values, expected by chance in a contingency table.

#### 4. PROPERTIES OF THE ELEMENTS

By now the student will very likely be speculating as to the nature of these elements. Their identification with physical entities seems unlikely. Psychologists working on mental tests where equiproportion is common have occasionally speculated on their being eventually identifiable with synapses, or genes, or stimuli, or unit experiences or other entities. It seems more probable that they are mathematical abstractions, units conceived of

\* For Eqs. 21-22, Ch. VI, see notes at end of chapter.

because they interpret data in orderly quantifiable ways and finding their justification in that fact.

Four properties of these elements are essential, namely, that they are *small, independent, frequency variables*. To say they are small entails the corollary that they are numerous, since their sum gives the observed indices. Their number may be of the order of thousands or millions to produce perfectly normal data, but a hundred or two have been shown to give normal results and equiproportions within probable error limits. To say that they are independent is to affirm absence of correlation between the elements. One could work with correlated elements, but the need would still exist to further analyze such elements until more ultimate uncorrelated elements were attained. This property meets the need in the social sciences for non-overlapping categories of thought. Finally, the elements are not constants but vary in an all-or-none or other fashion with some amount of frequency or probability of each value over which they vary. They thus have means, sigmas, and other properties of frequency distributions.

### III. TWENTY-FOUR SYSTEMATIC DIAGRAMS INTERPRETING CORRELATION

Since correlation is the general type of the relationship between quantitatively varying phenomena of which causation is that subtype where a time sequence is also involved, the student should be able to interpret correlation readily in all its different forms. To cultivate this ability twenty-four diagrams are presented below showing different ways of interpreting correlation from its computation to the properties of special correlation surfaces, or to functions of correlation coefficients. Each diagram (with one or two exceptions) shows positive and negative correlation and amounts of correlation in four steps, namely: zero, one quarter, one half, and perfect ( $r = 0, .50, .707, 1.00$ , or  $r^2 = 0, .25, .50$ , and  $1.00$ ). These amounts of correlation are illustrated in order to show the two limits of correlation, and what from differing viewpoints is midway between the limits.  $r = .5$  is usually thought of as midway, but for the fundamental aim of science to predict  $r = .707$ ,  $r^2 = .5$  is midway, for at this point what can be predicted equals what cannot be predicted. The known factors

in common to the two indices correlated equal their unknown factors of difference, and the coefficient of correlation equals the coefficient of alienation,  $r = k = .707$ . (See S. 12, Ch. VI.)

Throughout these systematic diagrams the notation is that of S-theory consistent with the rest of this volume. The more conventional X, Y, and Z, denoting variates, are our indices, I, J, and K, and the conventional N denoting the number of cases, i.e., the statistical population, becomes P here where we are limiting Sociology to human populations. The diagrams and formulae are general to biological, economic, or physical correlations (within limits of the assumptions and definitions) but are presented here for the student of Sociology, and for consistency the notation is standardized throughout this sociological system.<sup>30</sup>

To understand these diagrams calculus is not needed, but High School algebra, geometry, and trigonometry, plus a course in statistics are expected of the reader.

Following these twenty-four systematic diagrams are nineteen S-situations, sampled from the social science literature illustrating correlation in varying forms and illustrating S-notation for symbolizing correlation flexibly and compactly. For explanation of the "descriptive formula" and "legend" the paragraphs preceding the S-situations at the end of Chapter II may be consulted.

#### DEFINITIONS:

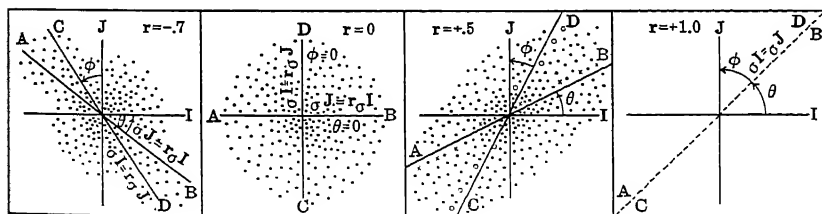
I, J, K = observed variables in units of deviation from the mean

$\sigma I = I/\sigma_I$  = standard score of I, or I in sigma units. Similarly  $\sigma J$ ,  $\sigma K$

P = the population

$r_{IJ} = \Sigma \sigma I \sigma J / P =$  the correlation coefficient of I and J,  $= \Sigma IJ / P \sigma_I \sigma_J$

#### S. 1, SCATTERGRAM + REGRESSION



"The smaller the angle between regression lines, the higher the correlation"

$\bar{x}$  = means of J arrays      0 = means of I arrays

$CD = [\sigma I_E = r \sigma J]$  regression of I upon J (I estimated knowing J) (Eq. 23, Ch. VI)

$AB = [J_E = r_I I]$  regression of  $J$  upon  $I$  (Eq. 24, Ch. VI)  
(estimated, i.e., most probable  $J$  for a given  $I$ )

$\tan \theta = \sigma_J / \sigma_I r$ ;  $\tan \phi = \sigma_I / \sigma_J r$  (Eq. 25a and b, Ch. VI) <sup>31</sup> \*

(Eq. 26, Ch. VI)

When  $\sigma_I = \sigma_J$ ,  $\tan \theta = \tan \phi = r$  (Eq. 25a, Ch. VI)

The fundamental meaning of correlation as a cross-classification of the two indices of two characteristics is revealed in the scattergram. A dot represents a person. Its co-ordinates represent the amount of each of the two indices characterizing that person. The class-intervals of one index are represented by the rows of a tabulation and the class-intervals of the other by the columns,  ${}_i(I) :: {}_j(I)$ . In each cell at the intersection of row and column there is a corresponding frequency of persons. This is symbolized in the basic matrix formula,  ${}_i(I) :: {}_j(I) : P$ . When the frequencies are stated as proportions of the whole population,  ${}_i P$ , they are *probabilities* of a person having jointly the amount of each characteristic denoted by that row and column.

The two regression lines, defined by Eqs. 23 and 24 for rectilinear regression, enable the estimating of the most probable amount of one index (within defined limits of error) from a knowledge of the other index (assuming a previously observed correlation in a persisting or comparable situation). If the unknown index occurs later in time, the estimate of it is a *prediction*. For a given amount of one index,  ${}_i(I)$ , the mean of the dots in the distribution of that column is the most probable amount of the other index,  ${}_i(I) : {}^M(J)$ . The regression line is taken as a convenient approximation to these means. The estimate may be in error as measured by the amount of the dispersion in the array, which is  $\sigma k$ . Since  $k$ , the coefficient of alienation (see S. 12, Ch. VI), varies oppositely to  $r$ , the higher the correlation, the smaller the error of estimate. Hence the accuracy of prediction improves with higher correlation.

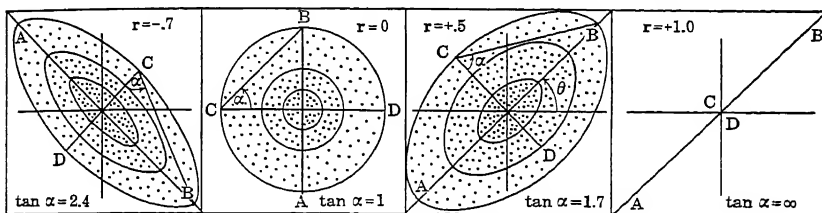
The slopes of the regression lines from their respective axes are given by Eqs. 25 and 26. When dispersions are equal ( $\sigma_I = \sigma_J$ ),  $r$  becomes the tangent of a regression line.

The scattergram is a powerful diagnostic tool. Inspection of it shows whether the data are markedly rectilinear or curvilinear, and therefore, what formulae will be most accurate to use. Nor-

\* For Eq. 25c, Ch. VI, see notes at end of chapter.

mality, unimodality, skewness, and other gross irregularities in the data can be visually spotted. The "falling down" phenomenon (see S. 29, Ch. VI) and all exceptional cases have the spotlight thrown upon them in the scattergram and can be identified for remedial treatment.

## S. 2, ELLIPSES (CONTOURS) + PRINCIPAL AXES



"The 'thinner' the ellipse, the higher the correlation"

$$\tan 2\theta = \frac{2r_{IJ}\sigma_I\sigma_J}{\sigma_I^2 - \sigma_J^2} \quad (\text{Eq. 27, Ch. VI})$$

$$\text{Family of ellipses} = \frac{I^2}{\sigma_I^2} - \frac{2r_{IJ}}{\sigma_I\sigma_J} + \frac{J^2}{\sigma_J^2} = d \quad (\text{Eq. 28, Ch. VI})$$

$d = \text{a constant for each ellipse}$

Taking  $\sigma_I = \sigma_J = 1$ :

$$\frac{AB}{CD} = \frac{\text{major axis}}{\text{minor axis}} = \frac{\sigma_{I+J}}{\sigma_{I-J}} = \frac{\sqrt{1+r}}{\sqrt{1-r}} = \tan \alpha \quad (\text{Eq. 29, Ch. VI})$$

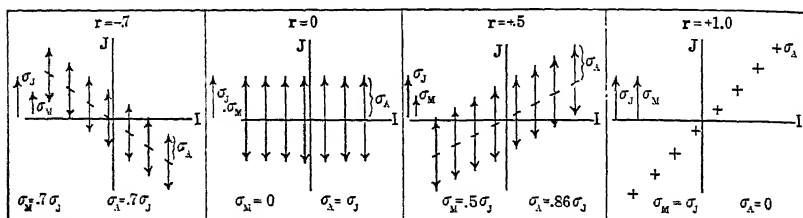
For negative correlation  $AB/CD = \sigma_{I-J}/\sigma_{I+J} = \sqrt{1-r}/\sqrt{1+r}$  in a diagram as drawn.

The curves are contour lines of the normal or bell-shaped correlation surface (see S. 5, Ch. VI). It summarizes the "scatteration" of points as an ellipse varying from a circle, when correlation is absent, to the other limit of a diagonal line when correlation is perfect. The slope of the major axis (which is at  $45^\circ$  when the sigmas are equal) is given by Eq. 27 (Ref. 34, p. 2). The family of ellipses may be plotted, one for any given value of the constant  $d$ , by assigning values for  $I$  and solving for the corresponding values of  $J$  in Eq. 28. The "fatness" of the ellipse is given by the ratio of its two axes which is conveniently specified as the tangent of the angle  $OCB$  as in Eq. 29.

These graphs illustrate, for the simple case of two indices, the analysis into the principal axes type of components developed by Hotelling and Kelley. The total variance of a set of indices is

transformed by this analysis into a maximal variance (the longest axis, AB) of one component: plus the next largest possible variance that is uncorrelated to the first, in a second component (the next to the largest axis, CD here), and so on for the succeeding components in as many further dimensions as there may be further intercorrelated indices in the recorded situation.

### S. 3, ARRAY VARIANCES + CURVILINEAR CORRELATION



“The smaller the dispersion in the array, the higher the correlation”

$\sigma_{MJ} = \sigma$  of means of the J arrays

$\sigma_{AJ} =$  weighted mean of sigmas of the J arrays

$\sigma_J =$  sigma of the distribution of J

*Rectilinear case:*

$$r = \sigma_{MJ} / \sigma_J \quad (\text{Eq. 30, Ch. VI}) \quad k = \sigma_{AJ} / \sigma_J \quad (\text{Eq. 31, Ch. VI})$$

$$\sigma_J^2 = \sigma_{MJ}^2 + \sigma_A^2 \quad (\text{Eq. 32, Ch. VI}) \quad l = r^2 + k^2 \quad (\text{Eq. 33, Ch. VI})$$

*Curvilinear case:*

$$\eta_{IJ} = \frac{\sigma_{MJ}}{\sigma_J} \quad (\text{Eq. 34, Ch. VI}) \quad \eta_{JI} = \frac{\sigma_{MJ}}{\sigma_I} \quad (\text{Eq. 35, Ch. VI})$$

The arrows represent the dispersion of the J values in each array from  $+\sigma$  to  $-\sigma$  on either side of the mean of the array, which is defined by a class-interval of the I indicant. The case of rectilinear regression is graphed, although this analysis by arrays is used more in calculating the two correlation ratios,  $\eta_{12}$  and  $\eta_{21}$ , measuring curvilinear correlation. The homoscedastic case, where all arrays have equal dispersions so that the sigma of any array equals the mean of the sigmas of the arrays, is here graphed. Each correlation ratio may be calculated either via:

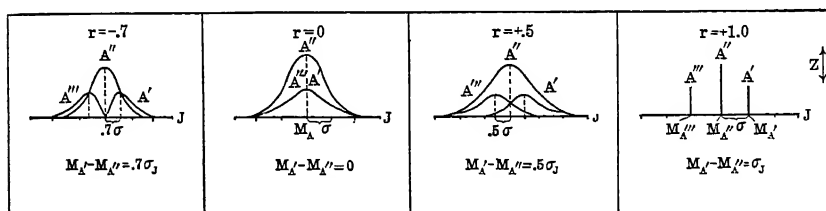
- the ratio of the sigma of the means of the arrays to the sigma of the distribution (Eq. 30, Ch. VI) (Ref. 22, #221), or
- the ratio of the mean of the sigmas of the arrays to the sigma of the distribution (Eq. 31, Ch. VI) (Ref. 22, #222). These

formulae are even more general than the Pearson product moment formula for rectilinear correlation, since Eqs. 34 and 35 include rectilinear and curvilinear correlation.

The correlation ratios (Eqs. 34 and 35) will be positive as computed. Whether the correlation is positive or negative must be determined by inspection of the data.

The variance of the arrays,  $\sigma_A$ , measures the part of the total variance,  $\sigma_J^2$ , of one index, (J), which is unknown or unpredictable from the other index (I) except as a probability. The variance of the means of the arrays,  $\sigma_M^2$ , measures the residual part of the total variance of (J) which is known and predictable from a knowledge of (I). The variance of the means plus the variance of the arrays comprise the total variance (Eq. 32). These two constituent variances are related as the coefficients of correlation and alienation (Eq. 33), or as cosine and sine (see S. 12, Ch. VI). The aim of better prediction in Sociology involves finding or inventing indices yielding smaller array array dispersions when correlated with the index to be predicted.

#### S. 4, OVERLAPPING DISTRIBUTIONS + BISERIAL CORRELATION



"The less the overlap, the higher the correlation"

3 of the I arrays, at  $-\sigma$ , 0, and  $+\sigma$  are graphed as frequency distributions of J values.  $M_{A'}$  = mean of an array

In the case of 2 arrays only:

$$\text{Biserial } r = \frac{(M_2 - M_1)pq}{\sigma z} \quad (\text{Eq. 36, Ch. VI})$$

p, q, z = normal probability table readings

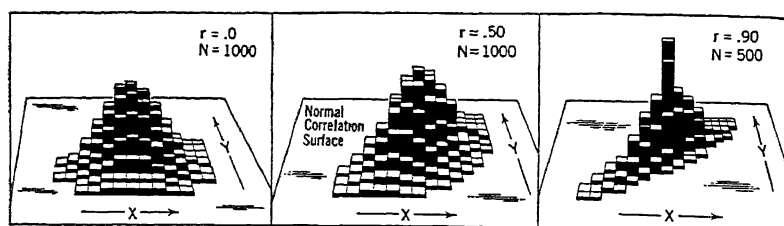
The overlapping distributions of the arrays of S. 3, Ch. VI are here viewed from another angle. If parallel vertical planes were passed through the correlation solid shown in S. 5, Ch. VI, their sections would be the distribution curves which are shown here as all projected onto one plane. These array distributions vary

from complete overlap in the absence of correlation up to no overlap in perfect correlation. In negative correlation, as in the leftmost diagram, the meaning of  $A'$  and  $A''$  is the opposite from the other three diagrams, i.e., the  $J$  scale is reversed.

When one index is given in only two class-intervals there are but two array distributions of the other index, i.e.,  $A' + A''$  only. This is the case for which the biserial  $r$  formula is adapted.

See S. 31, Ch. VI for an illustration of this analysis of correlation into overlapping distributions.

### S. 5, CORRELATION SURFACE



“The more ridgelike the dome, the higher the correlation”

Unit normal surface where  $P = 1 = \sigma_I = \sigma_J$

(Eq. 37, Ch. VI)

$$z = .15915k^{-1}e^{-E}$$

(Eq. 38, Ch. VI)

$$E = (.I^2 + .J^2 - 2r_{IJ}.I.J)/2k^2$$

(Eq. 39, Ch. VI)

$$= d/2k^2 \text{ (d as in S. 2, Ch. VI)}$$

(Eq. 39a, Ch. VI)

The normal correlation surface and the solid under it are here shown for indicants in 13 and 15 class-intervals. Horizontal planes passed through this solid yield the elliptical contour lines of S. 2, Ch. VI, while vertical planes perpendicular to one axis passed through this solid yield the overlapping distribution curves of S. 4, Ch. VI. A single plane perpendicular to one axis and one other plane perpendicular to the other axis divide the solid into four parts and create the situation dealt with by the tetrachoric  $r$ .

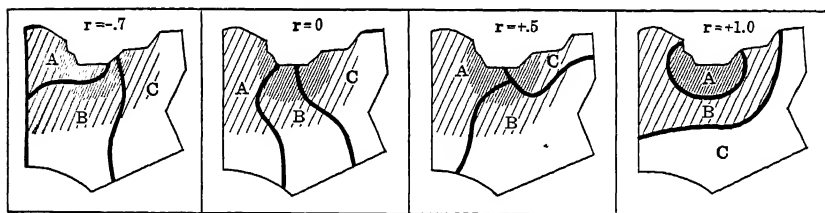
The first diagram for  $r = -.9$  is the photographic reverse of the fourth for  $r = +.9$  which was diagrammed instead of  $r = +1.0$  since the latter ceases to be a three dimensional surface as it shrinks to a distribution curve in a plane passed vertically through the main diagonal array.

In Eqs. 37 to 39  $z$  is any altitude from the base plane up to the surface,  $k$  is the coefficient of alienation,  $\sqrt{1 - r^2}$ ,  $e$  is 2.7183, the base of natural logarithms, and  $d$  is the equation of an ellipse

(Eq. 28, Ch. VI). (Ref. 35, #88.) The three surfaces of  $r = 0$ ,  $+0.5$ ,  $+0.9$  illustrate the situation in regard to prediction of socially controlled institutional behavior in the case of graduation from college. Graduation standing, ranging from failure-to-graduate up to "highest honors," shows zero correlation with irrelevant indices such as, the weight, or size of ears of the applicants to Freshman year. Graduation standing correlates around .50 (as a central tendency varying with the college) with a team of admission indicants such as, rank in class in secondary school, College Board examinations, aptitude tests, and chronological age. The "ceiling," or upper limit, that may reasonably be expected is a correlation of .90 between graduation standing and the standing of Freshman through Junior year.

In terms of the coefficient of determination ( $r^2$ ), expressed as a percentage, it may be said that size of ears yields 0% of prediction, a team of admission indices yields about 25% of perfect prediction, and the first three years of college yield about 80% of perfect prediction of graduation standing.

#### S. 6, SPATIAL CORRELATION-MAPS



"The more congruent the zones, the higher the correlation"

$$\sigma_I^2 = \Sigma I^2 / L^2 \quad (\text{Eq. 40, Ch. VI})$$

$$r_{IJ} = \Sigma I_J J / L^2 \quad (\text{Eq. 41, Ch. VI})$$

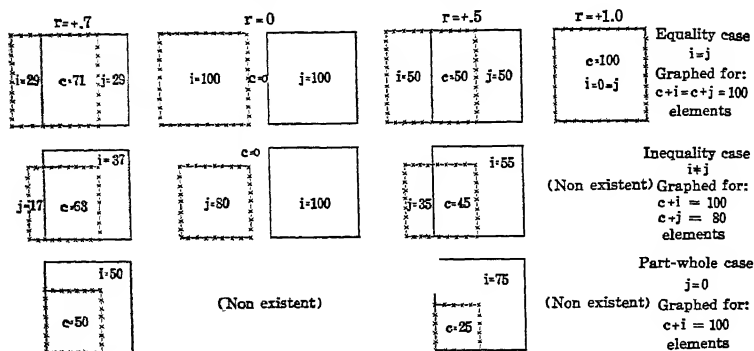
where  $L^2$  = number of districts, square miles, or other areal units

Example of a port city zoned by variable I in 3 class-intervals (shadings) vs. variable J in 3 classes, A, B, C.

Geographic correlation means the degree to which the zones of a region, classified on the basis of one characteristic, tend to coincide with the zones classified on the basis of a second characteristic. If the corresponding zones of the two sets are congruent, correlation is perfect between the two characteristics; if the corresponding zones cross-cut each other completely, correlation is

absent. When the characteristics are qualitative ones, their indicators are attributes and a contingency coefficient measures their relationship (see S. 2, Ch. VIII); when the characteristics are quantitative ones, their indicators are indicants and a correlation coefficient measures their relationship (see S. 21 and 22, Ch. VIII). The contingency coefficient also measures the relationship of a qualitative to a quantitative characteristic as in ecological studies determining the correlation of racial quarters of a city and its density zones, or of political districts and the density of some characteristic (see S. 19, Ch. VIII).

#### S. 7, PERCENT OF COMMON ELEMENTS OR OVERLAPPING MEMBERSHIP IN 2 GROUPS



"The larger the common factor, the higher the correlation"

$c$  = number of elements common to I + J

$i$  = number of elements specific to I

$j$  = number of elements specific to J

$$r_{IJ} = \sqrt{\frac{c}{c+i} \frac{c}{c+j}} = \text{geometric mean of 2 proportions} \quad (\text{Eq. 42, Ch. VI})$$

$$\text{when } j = 0, \quad r_{IJ}^2 = \frac{c}{c+i} \quad (\text{Eq. 43, Ch. VI})$$

$$\text{when } i = j, \quad r_{IJ} = \frac{c}{c+i} \quad (\text{Eq. 44, Ch. VI})$$

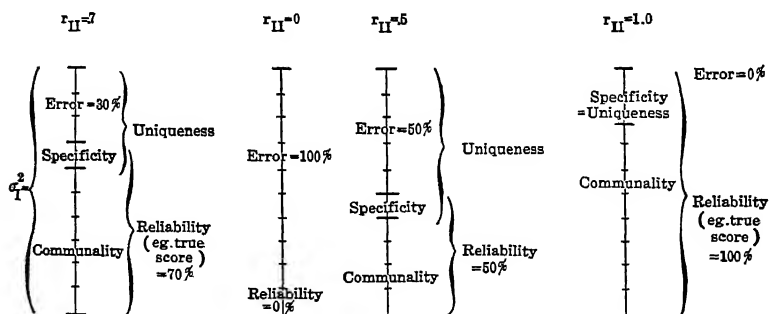
The formulae for correlation as the geometric mean of the two ratios (of the number of common elements divided by the total number of elements in each of the two indices correlated) has been most used in situations artificially synthesized by throwing groups of colored dice (Ref. 22, #297-9). In order to test hypotheses of the probable structures of components in indices,

especially in the field of intelligence measurement, such dice elements and the formulae above have been very useful. To analyze sociometric indices into such elements involves assumptions, the reasonableness of which is often unknown in our present stage of research. Such analyses, however, have possibilities of great usefulness which may abundantly reward research upon them.

For one example, the overlapping membership of persons in various groups (such as fraternal lodges) is one field of application of these common elements formulae which has been very little studied by sociologists as yet. Such persons are probably not elements which are independent of each other, as friendships will produce correlations in behavior. Such correlation between persons will complicate the analysis. (See S. 24, Ch. VI for a technic of determining the correlation between two persons.)

When one index is a part of the other, as an addend is to a sum, the common elements formula simplifies to Eq. 43, Ch. VI. This situation of a component, wholly and exclusively constituting a part of an index, reappears from differing viewpoints as a "cause" in Eq. 82, as "specificity" and "error" and "uniqueness" in Eqs. 45 and 47, as orthogonal axes in Eq. 27 and in Eqs. 72 and 77, and as mean array variance in Eqs. 31 and 32. At first glance these seem to have little in common, but they are all interconvertible mathematically and all express the relationship of a whole variable phenomenon to a part of it which is uncorrelated with the other parts.

### S. 8, RELIABILITY + ERROR VARIANCE



"The smaller the error, the higher the reliability correlation"

$I'$  = remeasurement of  $I$

$$\sigma_I^2 = \sigma_h^2 + \sigma_s^2 + \sigma_e^2 \quad (\text{Eq. 45, Ch. VI})$$

$h$  = communality

$s$  = specificity

$e$  = error

$$\sigma_e = \sigma_I \sqrt{1 - r_{II'}} = \text{error of measurement} \quad (\text{Eq. 46, Ch. VI})$$

Variance of true score =

$$\sigma_{I_{II'}}^2 = \sigma_h^2 + \sigma_s^2 = \sigma_I^2 - \sigma_e^2 \quad (\text{Eq. 47, Ch. VI})$$

Multiple independent errors =

$$\sigma_e^2 = \sigma_I^2[(1 - r_{II'}) + (1 - r_{II''}) + \dots + (1 - r_{II^*})] \quad (\text{Eq. 46a, Ch. VI})$$

$$r_{\infty\infty} = r_{IJ} / \sqrt{r_{II'} r_{JJ'}} \quad (\text{Eq. 48, Ch. VI})$$

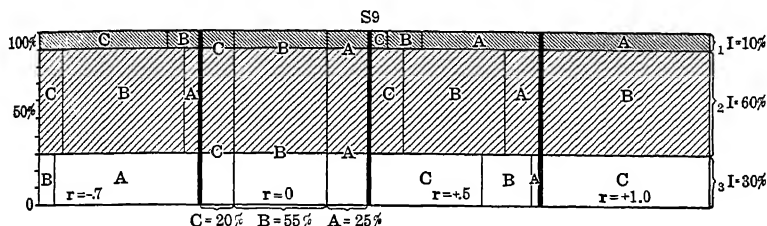
=  $r_{IJ}$  corrected for attenuation from errors

The variance of each index of a set observed in one population may be analyzed into three parts, namely, a subvariance due to errors, a non-error subvariance that is specific to that index alone, and a communality subvariance that is common to the other indices. (Ref. 77, p. 62.) The error and specificity together are called the "uniqueness" of that index. The specificity and the communality together are the true variance, which is also called the reliability, or the indicant free of errors.

The measurement and the analysis of the unreliability of observations is an essential step in making Sociology more of a science. Error can be surely reduced only when its amount and source are known. Thus, in a controlled experiment on rural hygiene (Ref. 12, pp. 62 ff.) the reliability of the hygiene scores observed among the families surveyed was measured by comparing means on remeasurement to discover the amount of any constant errors, and by correlating the two measurements to discover the amount of any variable errors. The measurements were experimentally made so as to analyze sources of error. By repeating the survey with differing interviewers, while all other factors were constant, the error due to the interviewers was isolated. By repeating the survey and obtaining the data from different informants in each family, with other conditions unchanged, the error due to the informant was isolated. Similarly the errors due to the schedule card, to the season of the year, and to the scorer of the schedules, were measured. By determining the intercorrelation of these errors, their synthesis of all errors was finally determined as 14% of the observed variance of hygiene. Such analyses had previously led to technics which had reduced the errors enormously, e.g., in the

case of scorer error from a 60% error ( $k^2$ ), as the schedule card was originally constructed, to a 0% error, as it was finally reconstructed.

### S. 9, PERCENTAGE BLOCKS



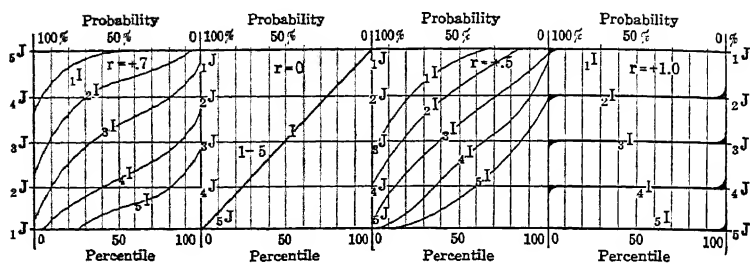
"The more congruent the zones, the higher the correlation"

Shaded strata = % of cases in each class of variable I; lettered zones running vertically = sub-percent in each class of J. (To get perfect correlation the percentages of A, B, and C must change to 10, 60, and 30 in the diagram for  $r = +1.0$ .)

Percents graphed are observational ones; theoretic ones as from a normal correlation surface are also possible. This graph also represents contingency between qualities.

The percentage blocks diagram is useful among laymen who shy off when correlation and prediction are presented in terms of a standard deviation of arrays or other statistical Greek. Homely percentages are readily interpreted. The strata may be the classes of the known or predictor variable (where one precedes the other in time), as in the standing of Freshman in unequal intervals of "With honor," "Passed," and "Failed." The graphed classes instead of being steps of a quantitative variable could equally well be qualitative categories. The diagram would then represent a contingency table. Each unit of area represents an absolute frequency of cases, as well as relative or percentage frequencies of the whole population and of its subclasses. The size of the correlation may be roughly gauged by the tendency of the steplike boundaries between subclasses of one kind to go from the vertical ( $r = 0$ ) to the horizontal ( $r$  approaches 1.0).

## S. 10, PROBABILITY WHORL



"The fatter the whorl, the higher the correlation"

Each curve represents a class-interval of predictor variable I. The co-ordinates of any point on a curve give the probability,  $\%P$ , of a case reaching or exceeding a given point on the predicted variable, J.

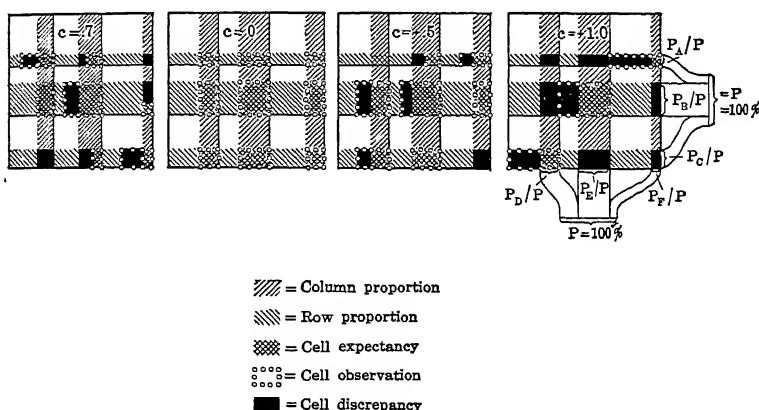
$$100 - \%P = \text{percentile}$$

This whorl of cumulative percentage curves is convenient for laymen in reading the probability of any amount of the J indicant for a given amount of the I indicant (given a correlation between the two in past experience which is assumed to continue to hold). Thus if John X is below average in being in the  $4I$  class-interval on a predictor indicant (such as the criteria for college admission added together with optimal weighting by multiple correlation technics), the  $4I$  curve is selected in the graph of  $r = .5$  (if .5 has been the correlation of such admission criteria with graduation standing). This curve is followed to its intersection with any level of graduation standing, J, which it is desired to predict, and the corresponding probability is read as the ordinate to the point of intersection. The probability for example of John X attaining a graduation of  $1J$  is zero, of attaining  $2J$  is 10%, of  $3J$  is 35%, etc. In percentile terms, if he should attain a  $2J$  standing later he would be a 90 percentile *in his admissions group of those in the  $4I$  class-interval*. The percentile rating is the complement of the probability. This is a valuable diagnostic tool for educational administrators. A percentile below 50 identifies students whose achievement is below expectation relative to their own previous record and ability. Laziness, disinterestedness, extra-curricular distractions, or similar causes calling for investigation and treatment towards motivating the student to greater academic effort are indicated, even though the student may be passing safely.

These probability curves are a refinement of achievement quotients, the ratio of educational test age and mental test age, which are used more at elementary school levels.

Curves for all values of  $r$  for indicants in sigma units could be computed from the normal correlation surface (Eq. 38, Ch. VI) and published so as to be available for routine predictions of this sort. With the use of "arithmetic probability paper" the ogives would become straight lines and a simpler set of nomographs could be constructed for prediction purposes.

### S. 11, CONTINGENCY



"The larger the cell discrepancies, the higher the contingency"

$,P$  = proportion of cases in any column (Eq. 49a, Ch. VI)

$,,P$  = proportion of cases in any row (Eq. 49b, Ch. VI)

$,,,P$  = proportion of cases in any cell (observed) (Eq. 49c, Ch. VI)

$,P,,P$  = proportion of cases in any cell (theoretic) (Eq. 49d, Ch. VI)

$,,,P - ,P,,P$  = cell discrepancy (Eq. 49e, Ch. VI)

$\Sigma ,P = \Sigma ,,,P = \Sigma ,,,P = 1.0$  (Eq. 49f, Ch. VI)

$$\phi^2 = \frac{\Sigma \{ (,,,P - ,P,,P)^2 \}}{,,,P} \quad (\text{Eq. 50, Ch. VI})$$

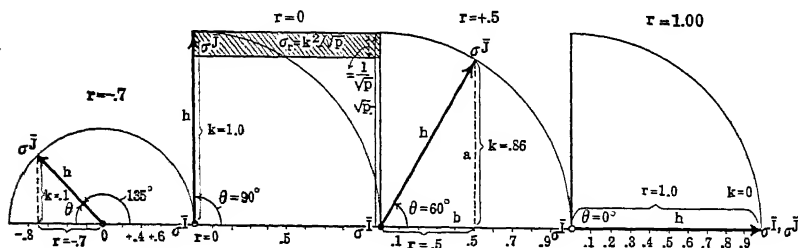
$$C = \sqrt{\phi^2 / (1 + \phi^2)} = \text{coefficient of contingency} \quad (\text{Eq. 51, Ch. VI})$$

=  $r$  for graduated "normal" data

The coefficient of contingency, when applied to quantitative, normally distributed data in many classes, becomes almost equivalent to the correlation coefficient. The case of two 3-category

attributes is here diagramed. In each of the four graphs, the height of the shaded area in each row is the probability for that row, and the breadth of the shaded area is the probability for that column. Their doubly shaded intersection in each cell is their product, i.e., the theoretic joint probability of a case being simultaneously in that row and column. The line of little circles shows the observed frequency and the black area shows the discrepancy from the theoretic frequency expected by the law of joint probability. The sum of these discrepancies squared gives Pearson's chi-square goodness-of-fit test,  $X^2$ , of the data to the hypothesis of chance distribution. This test reduced to a scale which runs from zero to nearly unity (depending on the number of class-intervals) is the coefficient of contingency.

### S. 12, COSINE AND SINE



"The smaller the angle, the higher the correlation"

$$r = \cos \theta = \frac{\text{base}}{\text{hypotenuse}} = \frac{b}{h} \quad (\text{Eq. 52, Ch. VI})$$

$$k = \sin \theta = \frac{\text{altitude}}{\text{hypotenuse}} = \frac{a}{h} \quad (\text{Eq. 53, Ch. VI})$$

in right triangle a, b, h (see when  $r = .5$ )  
in unit circle, i.e.,  $\sigma_I = \sigma_J = 1 = h$ :

$$r = \text{projection of } \sigma_J \text{ on } \sigma_I \quad (\text{Eq. 54, Ch. VI})$$

$$(\text{or of } \sigma_I \text{ on } \sigma_J) \quad (\text{Eq. 52a, Ch. VI})$$

$$1 = r^2 + k^2 \quad (\text{Eq. 33, Ch. VI})$$

$$k^2 = 1 - r^2 \quad (\text{Eq. 55, Ch. VI})$$

$$\sigma_r = k^2 / \sqrt{P} \quad (\text{Eq. 56, Ch. VI})$$

The fundamental trigonometric interpretation of the correlation coefficient as the cosine of an angle and of the coefficient of alienation,  $k$ , as the sine of that angle is portrayed in these dia-

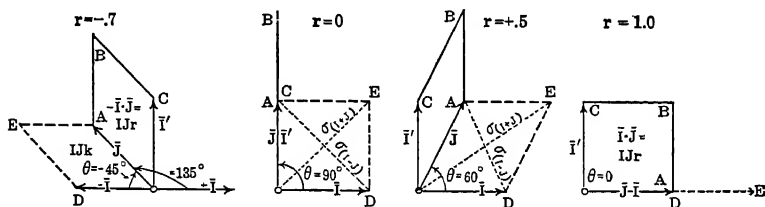
grams. A correlation of zero means a vectorial angle of  $90^\circ$ , a perfect correlation means a vectorial angle of  $0^\circ$ , i.e., completely collinear vectors ( $\bar{I} \equiv \bar{J}$ ).

The standard error of the correlation coefficient given in formula form in Eq. 56 is suggested as a relative area in the second graph showing  $r = 0$ .

The formula for the coefficient of alienation, Eq. 33, is seen to be a particular case of the familiar hypotenuse law (that the square of the hypotenuse equals the sum of the squares of the other two sides of a right triangle). The radius is taken as unity.

The interpretation of the correlation coefficient as a cosine is basic to the geometric representation of S-theory. Wherever correlations can be calculated, the angle between the two vectors representing the correlated indices becomes known. Amounts of any index determine lengths and specified points on its vector. With known lengths and angles, representing quantities and qualities respectively, societal n-space becomes determinate. It ceases to be a mere analogy and becomes subject to precise mathematical manipulation.

### S. 13, SCALAR PRODUCT OF 2 VECTORS



*Also area and diagonals of parallelogram*

“The more rectangular the parallelogram, the higher the correlation”

(Overlining denotes a vector)

$$\bar{I} \cdot \bar{J} = IJ \cos \theta = (OA)(OC) \cos \theta \quad (\text{Eq. 57, Ch. VI})$$

$$= \sigma_I \sigma_J r_{IJ} = p_{IJ} \quad (\text{Eq. 58, Ch. VI})$$

$$= \Sigma IJ/P = \text{covariance} \quad (\text{Eq. 59, Ch. VI})$$

With unit vectors, i.e.,  $I = J = \sigma_I = \sigma_J = 1$  (Eq. 60, Ch. VI)

$$\bar{I} \cdot \bar{J} = \cos \theta = r_{IJ} \quad (\text{Eq. 61, Ch. VI})$$

$$\bar{I} \cdot \bar{J} = \text{area } \square OABC \quad (\text{Eq. 62, Ch. VI})$$

formed by  $\bar{J}$  and  $\bar{I}'$ , the normal to  $\bar{I}$

$$\text{Area } \square \text{ ODEA} = IJ \sin \theta = IJk \quad (\text{Eq. 63, Ch. VI})$$

$$\text{Diagonals} = \bar{I} + \bar{J} = \text{OE} \quad \bar{I} - \bar{J} = \text{AD} \quad (\text{Eq. 64, Ch. VI})$$

$$\sigma^2_{(I+J)} = \sigma_I^2 + \sigma_J^2 + 2r_{IJ}\sigma_I\sigma_J = \text{OE} \quad (\text{Eq. 65, Ch. VI})$$

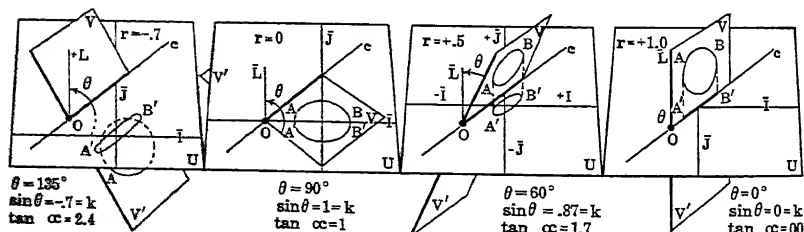
$$\sigma^2_{(I-J)} = \sigma_I^2 + \sigma_J^2 - 2r_{IJ}\sigma_I\sigma_J = \text{AD} \quad (\text{Eq. 66, Ch. VI})$$

The scalar product of two vectors, which is also the covariance of the two variables represented by the vectors, can be graphed as the area of the parallelogram formed by one vector and the normal (perpendicular) to the other. This parallelogram varies from zero area when it flattens out into a line (when  $r = 0$ ), to its maximal area when it fills out into a rectangle (when  $r = 1.0$ ). At this upper limit the scalar product becomes identical with the ordinary product of two lengths in everyday arithmetic. This is why in S-theory in determining the quantic, the scalar product is used, as it is the more general case with the arithmetic product as a special subvariety.

A most useful further fact is that the two diagonals of the other parallelogram (ODEA in S. 13), which are given in vector algebra as the sum and the difference of the two vectors, represent in statistical notation the standard deviation of the sum and the standard deviation of the difference of the two variables, respectively.

The two diagonals of the parallelogram formed by the two vectors are also the major and minor axes of the chief ellipse in S. 2, Ch. VI. The area of this parallelogram is  $\sigma_I\sigma_Jk$ , a kind of "complement" of the scalar product. (For these formulae for vectors and parallelograms see Ref. 32, pp. 15 and 64.)

S. 14, VECTOR PRODUCT OF 2 VECTORS



"The more upright the scatter-circle (AB), the higher the correlation," or,  
 "The 'thinner' the projected ellipse (A'B'), the higher the correlation"  
 (cf. #2)

$$\mathbf{I} \times \mathbf{J} = \bar{\mathbf{L}} I J \sin \theta \text{ (note right-handed screw rule for } \mathbf{L} \text{)} \quad (\text{Eq. 67, Ch. VI})$$

With unit vectors where  $\bar{\mathbf{L}} \perp$  plane U

$$\mathbf{I} = \mathbf{J} = \sigma_I = \sigma_J = \bar{\mathbf{L}} = 1 \quad (\text{Eq. 68, Ch. VI})$$

$$\bar{\mathbf{I}} \times \bar{\mathbf{J}} = \bar{\mathbf{L}} k_{IJ} \quad (\text{Eq. 69, Ch. VI})$$

$$\angle \theta = \angle \text{VOL} = 90^\circ - \angle \text{UOV} \quad (\text{Eq. 70, Ch. VI})$$

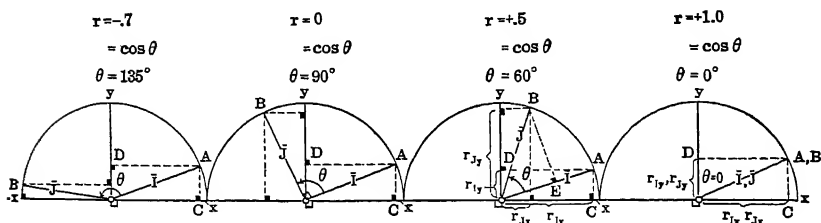
$$r = \cos \theta \quad (\text{Eq. 71, Ch. VI})$$

(See #2 for  $\tan \alpha$ )

The vector product of two vectors is a vector ( $\bar{\mathbf{L}}$  in S. 14) perpendicular to the plane (U) determined by the two vectors that are multiplied together ( $\bar{\mathbf{I}}$  and  $\bar{\mathbf{J}}$ ) (Ref. 32, p. 64). In terms of this vector-product the way can be graphed in which the contour ellipses of S. 2, Ch. VI may be generated. Let a circle, AB, in plane V be projected upon plane U giving the ellipse A'B'. Let the plane V rotate about the principal axis of the ellipse (or line parallel to it, such as OC). The projected ellipse will vary from a circle at one limit when plane U and V are parallel, so that the cosine of the angle between plane V and line L is zero ( $r = 0$ ) all the way to a line at the other limit when plane V is perpendicular to plane U, so that the cosine of the angle between plane V and line L is unity ( $r = 1.0$ ).

No important sociological use has as yet been discovered for the vector product of vectors in S-theory—in contrast to the abundant use of the scalar product of vectors. If the two vectors are unit vectors, their vector product is a vector whose scalar length is the coefficient of alienation,  $k$  ( $= \sin \theta$ ).

### S. 15, DIRECTION COSINES



“Any correlated variables are resolvable into uncorrelated components”

$$r = \cos, r_{IJ} = \cos \theta$$

The case of 2-axes is graphed.

For the general case of  $n$ -axes:

$$r_{IJ} = r_{Ix}r_{Jx} + r_{Iy}r_{Jy} + \dots + r_{In}r_{Jn} \quad (\text{Eq. 72, Ch. VI})$$

$$\sigma_I^2 = \sigma_{I\cdot x}^2 + \sigma_{I\cdot y}^2 + \dots + \sigma_{I\cdot n}^2 \quad (\text{Eq. 73, Ch. VI})$$

normalized (i.e., divided by  $\sigma_I^2$  as unit) gives:

$$1 = \sigma_{I\cdot x}^2/\sigma_I^2 + \sigma_{I\cdot y}^2/\sigma_I^2 + \dots + \sigma_{I\cdot n}^2/\sigma_I^2 \quad (\text{Eq. 74, Ch. VI})$$

$$= r_{Ix}^2 + r_{Iy}^2 + \dots + r_{In}^2 \quad (\text{Eq. 75, Ch. VI})$$

and similarly:

$$1 = r_{Jx}^2 + r_{Jy}^2 + \dots + r_{Jn}^2 \quad (\text{Eq. 76, Ch. VI})$$

$$r_{Iy} = k_{Iy}, r_{Jy} = k_{Jy}, \text{ etc.} \quad (\text{Eq. 77, Ch. VI})$$

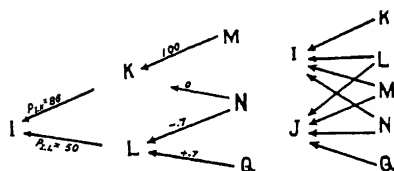
Some of the general properties of a direction cosine system are brought out in Eqs. 72 to 77, though the graphs show only the case of two variables in a two-dimensional space. A direction cosine is here the cosine of the angle between a reference co-ordinate and the vector. This is the sine,  $k$ , of the angle between the co-ordinate and the normal to the vector. The reference system of co-ordinates ( $C$ 's in S. 21, Ch. VI) is orthogonal, that is, the co-ordinates ( $x, y, z$ , etc.) are all mutually perpendicular.

$n$  oblique vectors can always be re-expressed in terms of  $n$  orthogonal axes, such as  $x, y, z$  in S. 15, which are mutually perpendicular reference vectors. Each vector such as  $\bar{I}$  is resolved into its component-vectors, which are its projections ( $OC$  and  $OD$ ) on the axes. The length of each projection is given by the square of the correlation between the unit vector ( $\bar{I}$ ) and the axis ( $r_{Ix}^2$ , etc.). The sum of the squares is the square of the unit vector  $\bar{I}$ . This is the hypotenuse law again, generalized to  $n$  dimensions (Eqs. 75 and 76). (See Ref. 32, p. 29.)

Sociologically the axes represent uncorrelated components in terms of which the correlated observed indices ( $I, J$ ) are re-expressed. Among other advantages, this fulfills the canon of classification by expressing phenomena in non-overlapping categories. Much confusion of thinking in the social sciences is due to using concepts which, being different words and not obviously synonyms, seem to be independent entities but represent statistically correlated phenomena and hence are overlapping categories.

## S. 16, PATH COEFFICIENTS, P

(Causal Analysis)



"The larger the path coefficient, the more complete the causation"

$$P_{I \cdot K} = \sigma_{I \cdot K} / \sigma_I \quad (\text{Eq. 78, Ch. VI})$$

$$\sigma_{I \cdot K}^2 = \text{variance of } I \text{ due to } K \quad (\text{Eq. 79, Ch. VI})$$

$$r_{IJ} = \sum_{K=1}^{K=Q} P_{I \cdot K} P_{J \cdot K} \quad (\text{Eq. 80, Ch. VI})$$

I, J = 2 effects

K, L, ..., Q = causes

For uncorrelated causes:

$$r_{IJ} = \sum_{K=1}^{K=Q} r_{IK} r_{JK} \quad (= (\text{Eq. 72})) \quad (\text{Eq. 81, Ch. VI})$$

For multiple causes:

$$100 = 100 \sum_{K=1}^{K=Q} \beta_{IK \cdot LM \dots Q}^2 + 100 \sum_{L=1}^{L=Q} \beta_{IK \cdot LM \dots Q} \beta_{IL \cdot KM \dots Q} r_{KL} \quad (\text{Eq. 82, Ch. VI})$$

$$\beta_{IK \cdot LM \dots Q} = P_{IK \cdot LM \dots Q} = \frac{(-1)^{K-1} \Delta_{IK}}{\Delta_{II}} \quad (\text{Eq. 83, Ch. VI})$$

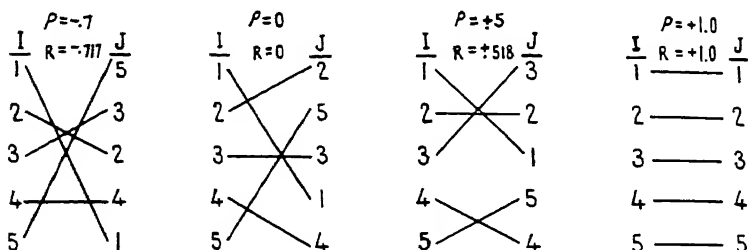
$$100 \beta_{IK \cdot LM \dots Q}^2 = \% \text{ of } \sigma_I^2 \text{ directly caused by } K \quad (\text{Eq. 84, Ch. VI})$$

$$100 \beta_{IK \cdot LM \dots Q} \beta_{IL \cdot KM \dots Q} r_{KL} = \% \text{ of } \sigma_I^2 \text{ jointly caused by } K + L \quad (\text{Eq. 85, Ch. VI})$$

The arrows represent the directions of causal "paths." The strength of the partial causes of a phenomenon are measured by functions of partial variances and regression weights ( $\beta$ ). For some situations, involving time sequence, cause and effect relationships may be known, or reasonably inferable, and their relative amounts may be measured by path coefficients. Thus the causal contribution of rainfall, acreage planted, and other factors to the final crop harvested can be determined. The coefficient of non-determination,  $k^2$ , is here again invaluable as it measures the residual unknown causation, challenging to further research.

In most social situations, however, the variables observed interact as cause and effect of each other or of a still unmeasured complex of further variables. Here causal inferences from correlation coefficients alone is illegitimate and the technics of path coefficients are inapplicable. (For fuller discussion and bibliography see Ref. 24.)

### S. 17, RERANKING



"The more horizontal the rungs of the ladder, the higher the positive correlation"

$$\rho = 1 - \frac{6\Sigma(I - J)^2}{P(P^2 - 1)} \quad (\text{Eq. 86, Ch. VI})$$

P = number of ranks in I (or J) (P = 5 is graphed)

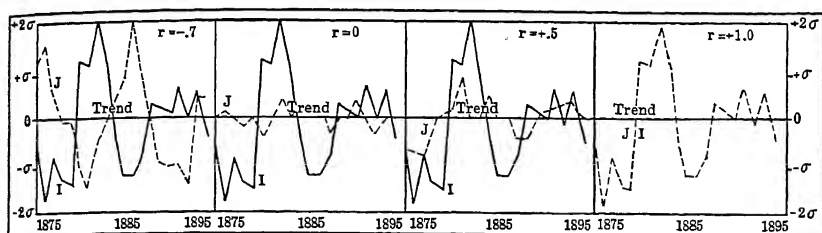
$$r = 2 \sin(\pi/6\rho) \quad (\text{Eq. 87, Ch. VI})$$

$$\sigma_1 = \sqrt{\frac{P^2 - 1}{12}} \quad (\text{Eq. 88, Ch. VI})$$

$$= .29P \text{ for } P \text{ large} \quad (\text{Eq. 88a, Ch. VI})$$

Correlation of ranks is a familiar form of correlation where the observed indicants are in ordinal units ( $^oI$ ) and have not been reduced to the more accurate cardinal units ( $^cI$ ) (multiples of equal, standardized units). If the underlying ranked data are normally distributed, the small correction (Eq. 87) for the rectangular distribution necessitated by ranking converts the rank correlation coefficient, rho ( $\rho$ ), into the Pearson product moment,  $r$ . In negative correlation, if one scale were reversed, the horizontalness of the "rungs" of the diagrammatic ladder would also increase as the correlation approached minus one. For formulae see Ref. 35, p. 193. Eq. 86 is that special case of Eq. 99, Ch. VI which results when the two equal sigmas, given by Eq. 88, are substituted into Eq. 99.

## S. 18, TIME SERIES



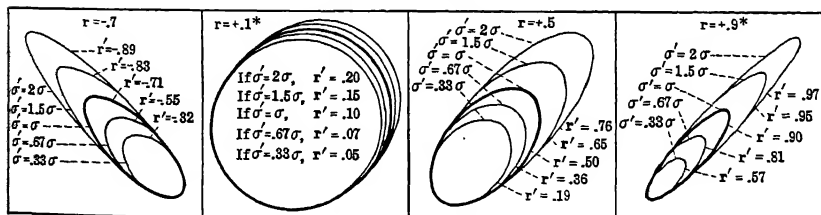
*Cyclic correlation with secular trend, seasonal variation, and lag removed*  
 "The more parallel the curves, the higher the positive correlation"

$$r = \frac{\sum_{I=1}^T (I - \bar{M}I) \frac{1}{\sigma} (J - \bar{M}J)}{t} \quad (\text{Eq. 89, Ch. VI})$$

where:  $t$  = number of annual dates, or periods, observed  
 $T$  denotes an  $I$  on one date  
 $L$  denotes a  $J$  on one date, such that  
 $L - T$  = lag interval (Eq. 90, Ch. VI)  
 $\bar{M}I, \bar{M}J$  = trend readings  
 $\sigma$  denotes "in sigma units"

Correlation of time series is possible in various ways. The graph here, S. 18, shows the cyclic relation of two indices measured in annual deviations from trend in sigma units. Seasonal fluctuations and the influence of long-time trend are thus eliminated. The experimentally determined lag may be allowed for and its obscuring effect removed as indicated in Eq. 89 and 90, Ch. VI. The graph of  $r = -.70$  is the slightly modified correlation (actually  $r = -.67$ ) between wholesale prices indicating business prosperity and church attendance in the United States. (Ref. 57, p. 169.) The others are of those prices and a hypothetical second variable yielding the correlations graphed. The formula is the standard product moment formula for time series in S-notation.

## S. 19, RANGE



\* Note.—Altering the range does not alter  $r$  at the limits of 0 or  $\pm 1.0$ .  
 $\therefore r = .1$  and  $.9$  are graphed instead.

"The longer the range, the higher the correlation"

Let  $\sigma$  = observed range  
 Let  $\sigma'$  = altered range  
 Let  $r$  ( $= \cos \theta$ ) be from  $\sigma$   
 Let  $r'$  ( $= \cos \theta'$ ) be from  $\sigma'$

$$\sigma/\sigma' \tan \theta = \tan \theta' \quad (\text{Eq. 91, Ch. VI})$$

or

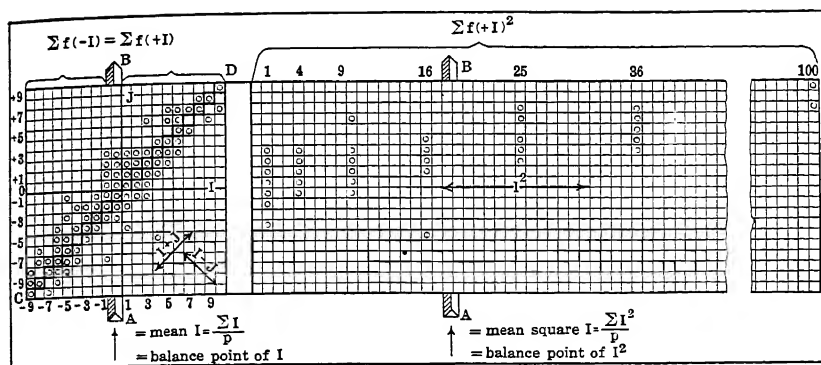
$$\frac{\sigma k}{\sigma' r} = \frac{k}{r'} \quad (\text{Eq. 92, Ch. VI})$$

Ellipses as in #2

The diagram, S. 19, attempts to dramatize the important fact, little appreciated by people who interpret correlations with a superficial knowledge of correlation theory, that the size of a correlation coefficient is partly dependent on the range in which it is observed. The correlation coefficient is expressed in sigma units so that if a population of greater dispersion is observed, the  $r$  will be larger; while if a population of curtailed dispersion is observed, the  $r$  will shrink. A given correlation between two tests in a population of one school grade can be raised to a higher correlation by including other grades, and vice versa. For each of the four values of  $r = -.7, +.1, +.5$ , and  $+.9$  the ellipse of the scattergram is shown with its rounding out as the range shrinks one third and two thirds; and its flattening in as the range expands by 50% and 100%. Thus a correlation of .50 can vary from .19 up to .76 as the range varies from one third up to twice the original range. This cautions the student to compare correlation coefficients of different populations in the light of the comparability of their ranges. For this the sigmas of the two variables should always be reported. Also, in addition to the raw correlations, those coefficients converted so as to be of a standard, or common, range should be reported, if the data are comparing different ranges.

The formula (Ref. 22, #295) for correcting for range can be very conveniently used in the form of Eq. 91 by reading the given cosines and corresponding tangents and the new cosines from a table of the natural trigonometric functions.

## S. 20, MOMENTS



"The larger the cross-product moment, the higher the correlation"

1st moment (origin = mean)

$$\Sigma I/P = 0 = \Sigma J/P$$

(Eq. 93, Ch. VI)

$$\Sigma I'/P = \text{mean } I \text{ (origin} = 0)$$

(Eq. 94, Ch. VI)

2d moment (origin = mean)

$$\sigma_I^2 = \Sigma I^2/P$$

(Eq. 95, Ch. VI)

$$\sigma_I^2 = \Sigma I'^2/P - M^2 \text{ (origin} = 0)$$

(Eq. 96, Ch. VI)

$$\sigma_{(I-J)}^2 = \frac{\Sigma (I - J)^2}{P} = \frac{\Sigma I^2}{P} + \frac{\Sigma J^2}{P} - \frac{2\Sigma IJ}{P}$$

(Eq. 97, Ch. VI)

$$P_{IJ} = \frac{\Sigma IJ}{P} = \frac{\sigma_I^2 + \sigma_J^2 - \sigma_{(I-J)}^2}{2}$$

(Eq. 98, Ch. VI)

$$r = \frac{\Sigma IJ}{P\sigma_I\sigma_J} = \frac{\sigma_I^2 + \sigma_J^2 - \sigma_{(I-J)}^2}{2\sigma_I\sigma_J}$$

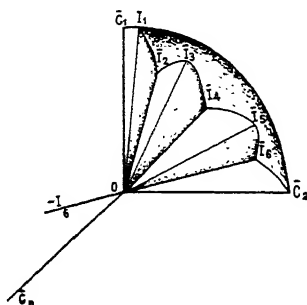
(Eq. 99, Ch. VI)

The relation of the statistical moments to the angular moments of force of Physics is here utilized in a mechanical device to calculate correlation by a gravimetric process. Each circle in the scattergram above represents a leaden weight symbolizing a person. The point of balance of the scattergram on the  $I$  axis is the mean  $I$ . On spreading the arrays apart as the squares of their deviations from the mean (as shown at the right), the mean square deviation is found as the point of balance. By spreading the diagonal arrays as the squares of the  $I - J$  deviations from the main diagonal array, the mean of the squared differences is determined by weighing and finding the point of balance. These quantities give  $r$  by the difference formula as in Eqs. 96-99, as well as both of the means and sigmas. (See formulae #79, 84, 85, 89, 156, 186 in Ref. 22.)

This relation of the statistical to the physical moments can be

utilized in a variety of mechanical calculating devices from the gravimetric method above to volumetric devices and to gear wheel and counter devices.<sup>32</sup>

### S. 21, INTERCORRELATIONS AND COMPONENT ANALYSIS



"Intercorrelations define a hyperspace of vectors and components"

$$\begin{array}{c|ccc|ccc}
 & I_1 & I_2 & \cdots & I_I & \cdots & I_n \\
 I_1 & V_1 & r_{12} & \cdots & r_{1I} & \cdots & r_{1n} \\
 I_2 & r_{21} & V_2 & \cdots & r_{2I} & \cdots & r_{2n} \\
 \hline
 I_I & r_{I1} & r_{I2} & \cdots & V_I & \cdots & r_{In} \\
 \hline
 I_n & r_{n1} & r_{n2} & \cdots & r_{nI} & \cdots & V_n
 \end{array} \quad \Bigg\| = R_{n,v} \text{ (Eq. 100, Ch. VI)}$$

$I$  = observed variable

$C$  = Cartesian co-ordinates, reference axes, orthogonal dimensions, or uncorrelated components

$r$  = cosine of angle between any 2  $I$ -vectors,  $\bar{I}$

$n$  = number of variables and order of matrix,  $\bar{I}$

$\Delta$  = determinant of  $R$

$n'$  = rank of  $R$  =  $n - m$

For components:

in variance units,  $v = 1$

(Eq. 101, Ch. VI)

in raw score units,  $v = \sigma^2$ ,  $r_{IJ} = p_{IJ}$

(Eq. 102, Ch. VI)

in true score units (reliability),  $v = r_{II}$

(Eq. 103, Ch. VI)

in communality units,  $v = h^2$  (cf. #8)

(Eq. 104, Ch. VI)

If every  $\Delta_{n'+1} = 0$  ( $\neq \sigma_\Delta$ )

(Eq. 105, Ch. VI)

the  $n$  variables lie in  $n'$  dimensions,

i.e., need  $n'$  components only

When  $m = n - 1 = 1$ ,  $\Delta_2$  = tetrad

(Eq. 106, Ch. VI)

and  $v = h^2$

(Eq. 107, Ch. VI)

the communality becomes one-dimensional,

i.e., 1 general +  $n$  specific components suffice

Let  $S$  = order of component, i.e., a component appearing in  $S$  variables, then components *must* exist

of order  $S$  (or higher) if mean  $r > (S - 2)/(n - 1)$  (Eq. 108, Ch. VI)  
 of order  $S - 2$  (or lower) if mean  $r < (S - 2)/(n - 1)$   
 (Eq. 109, Ch. VI)

$n$  observed indices ( $I$ ) may be represented by a sheaf of  $n$  vectors in  $n$ -dimensional space. Their intercorrelation coefficients form a square matrix ( $R_{n,r}$ ) of order  $n$ . These  $n$  indices may be re-expressed in terms of a smaller number,  $n'$ , of uncorrelated components (i.e.,  $n'$  Cartesian co-ordinates,  $\bar{C}$ ), whenever the matrix is of rank  $n'$ . One test of this rank is Eq. 105. The units in which the components are computed depend upon what values are chosen for the  $v$ 's as in Eqs. 101-104.

The purpose of such analysis of observed indices into mathematically calculated components is:

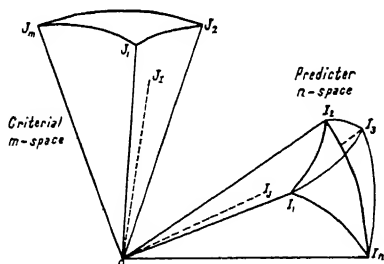
- a. parsimony—to express  $n$  indices in fewer components, and
- b. elimination of overlap—to describe data in terms of non-overlapping categories, i.e., in concepts that are uncorrelated with each other.

Geometrically this analysis represents a transformation of a set of oblique vectors ( $\bar{I}_n$ ) into a set of orthogonal vectors ( $\bar{C}_n$ ). The five changes noted above, which may be involved in such a transformation, are:

1. translation—shifting of the vectors, keeping each parallel to itself, till their origins coincide at 0, the apex of the sheaf of vectors
2. stretching—changing the length, the scalar units, of a vector
3. reflection—using the negative extension of a vector beyond its 0 point
4. rotation—rotating the sheaf to reorient it without distortion
5. distortion—altering the angles between vectors, i.e., falsifying the observed facts of intercorrelation.

Transformations may legitimately involve the first four in specified ways but the fifth is to be avoided (by preserving unchanged the evaluated determinant of the intercorrelations). (For Eq. 105 see Ref. 17, p. 275; for Eqs. 108 and 109 see Ref. 74, p. 158; for Eqs. 100-104, 106, and 107 see definitions in Ref. 77.)

## S. 22, MULTIPLE CORRELATION



"The smaller the minimal angle (between criterial and predictor spaces), the higher the multiple correlation"

One-way multiple R:

$$\begin{aligned} \text{when } \angle I_J O J_1 \text{ is minimal } (I_J \text{ is in } n\text{-space}) & \quad (\text{Eq. 110, Ch. VI}) \\ R_{J_1 \cdot I_1 I_2 \dots I_n} (= R_{J_1 \cdot I}) = \cos \angle I_J O J_1 & \quad (\text{Eq. 111, Ch. VI}) \\ = \sqrt{1 - \Delta / \Delta_{J_1 I}} \quad (v = +1, \text{ see S. 21 and Eq. 101}) & \quad (\text{Eq. 112, Ch. VI}) \\ = \sqrt{\sum \beta_{J_1 I_i \cdot I_1 \dots I_n}^2} & \quad (\text{Eq. 113, Ch. VI}) \end{aligned}$$

where  $i = 1, 2, 3, \dots, n$  in turn and  $i -$  denotes the  $i$  variables with the particular variable denoted by the prime omitted.

Two-way multiple R, "the most predictable criterion":

$$\begin{aligned} \text{when } \angle I_J O J_1 \text{ is minimal } (J_1 \text{ is in } m\text{-space}) & \quad (\text{Eq. 110a, Ch. VI}) \\ = \cos \angle I_J O J_1 & \quad (\text{Eq. 111a, Ch. VI}) \\ R_{I \cdot J} = \sqrt{r^{ih} r_{i\alpha} r_{h\beta} \alpha^\alpha \beta^\beta} & \quad (\text{Eq. 114, Ch. VI}) \end{aligned}$$

where:

1.  $r^{ih}$  denotes cofactors in  $\Delta$
2.  $i, h$  denotes each predictor in turn
3.  $\alpha, \beta$  denotes each criterion in turn
4. repeated scripts denote a summation

Multiple correlation is the largest possible correlation coefficient between a single dependent, or criterion, index and a team of optimally-weighted independent, or predictor, indices. Geometrically it is the cosine of the smallest possible angle between the criterion vector ( $\bar{J}_1$ ) and the predictor space, determined by the  $n$  predictor vectors ( $\bar{I}$ ).

The problem of prediction of societal phenomena is hereby reduced to a problem of finding, or inventing, measures whose vectors show:

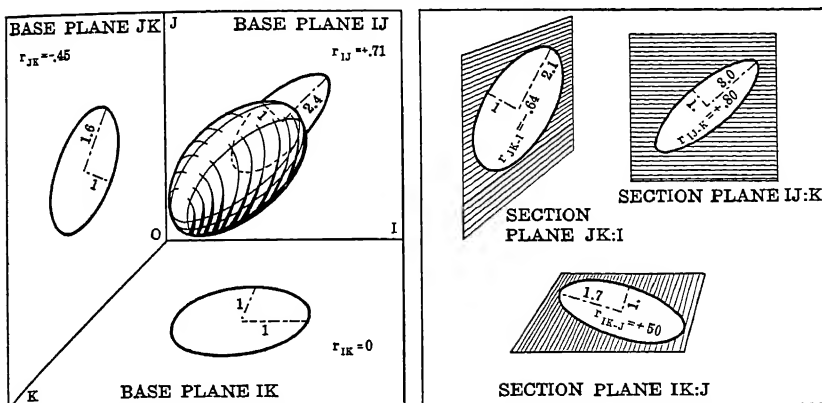
- a. smaller angles with the criterion vector, and
- b. larger angles with the other predictor vectors (as then the predictive contributions of the predictors overlap least).

Usually in science the chief problem is to obtain better data. The technics of multiple correlation merely make the prediction the best that is possible with given data.

An extension of multiple correlation is Hotelling's "most predictable criterion" (Ref. 30), which is a double multiple correlation to find the minimal angle between the vectors of a predictor team of indices and a criterion *team* of indices. The criterion characteristic may be a complex phenomenon, such as "community health," which is only partially indicated by an index of morbidity, or of mortality, or of longevity, or of sanitation, etc. But the composite of such indices might measure community

health more completely. The problem of improving community health then becomes one of developing those hygienic characteristics whose indices are found to give the highest double multiple correlation with the team of health criteria.

### S. 23, PARTIAL CORRELATION



"The sectional ellipse is to the projected ellipse as the partial to the total correlation"

$$r_{IJ:K} = \frac{r_{IJ} - r_{IK}r_{JK}}{k_{IK}k_{JK}} \quad (\text{Eq. 115, Ch. VI})$$

$$r_{IJ:KL} \dots N = \Delta_{IJ} / \sqrt{\Delta_{II}\Delta_{JJ}} \quad (\text{Eq. 116, Ch. VI})$$

$$= \sqrt{\beta_{IJ:KL} \dots N \beta_{JI:KL} \dots N} \quad (\text{Eq. 117, Ch. VI})$$

$$\beta_{JI:KL} \dots N = \frac{(-1^I)\Delta_{JI}}{\Delta_{JJ}} \quad (\text{Eq. 118, Ch. VI})$$

One geometric way of diagraming the first-order partial correlation between two indices with a third (K in S. 23 above) partialled out is to show it as a plane section of a three-dimensional ellipsoid. Let the scatter of points in S. 1 and S. 2 be extended into the third dimension to represent a third variable, K. Let this "swarm" of points be represented by contour ellipses as in S. 2, forming the ellipsoid floating in space in the diagram above at the left. If the three observed variables I, J, and K are graphed as Cartesian co-ordinates determining base planes IJ, JK, and IK, then the ellipse (on each of the base planes) formed as a projection or shadow of the ellipsoid represents the total (i.e., raw or zero order) correlation coefficient ( $r_{IJ}$ ,  $r_{JK}$ ,  $r_{IK}$ ). The ratio of the major to minor axes diagramed is given by Eq. 29.

If now a sectioning plane such as  $IJ : K$ , which is parallel to plane  $IJ$  and passes through the  $K$  axis at some particular point (such as the mean), is passed through the ellipsoid, the ellipse formed by the intersection of this plane and the ellipsoid represents the partial correlation of  $I$  and  $J$  for a constant value of  $K$  (i.e., with  $K$  "eliminated" or "partialed out"). The partial correlation is more exactly the average of such values, or sectional ellipses, only one of which is diagrammed here. (For the homoscedastic case any ellipse represents all the others in parallel planes, as the sigmas of arrays are all equal.)

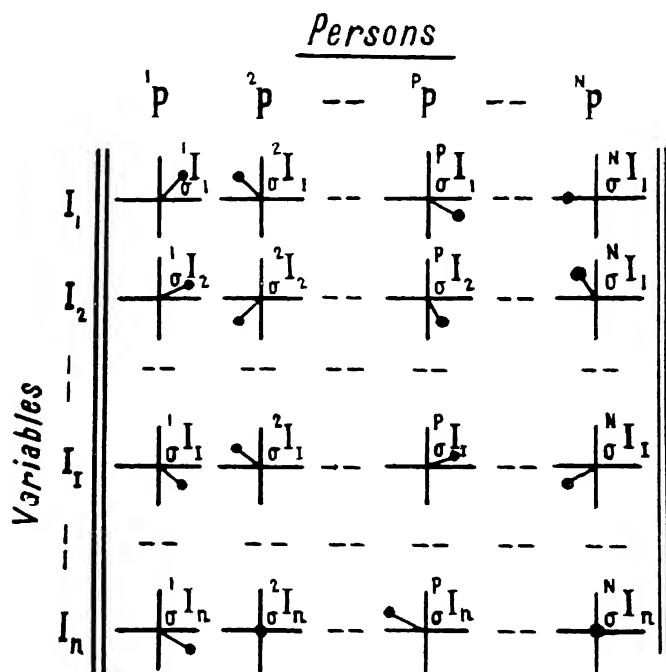
According as the three axes of the ellipsoid vary in length, and according as the orientation of the ellipsoid varies, the partial correlations may be greater or smaller, or even change in sign, compared to the raw correlations. A visually minded student may take a lump of plastecine and experiment with shadows and sections varying the axes so that the ellipsoid goes from one limit of a straight line to a planar circle and on to a cigar shape or to a sphere. This exercise may help some students to understand how a raw negative correlation can become a greater negative partial correlation (e.g.,  $r_{JK} = -.45$ ,  $r_{JK \cdot I} = -.64$  above), while a zero correlation jumps to a positive one (e.g.,  $r_{IK} = 0$ ,  $r_{IK \cdot J} = +.50$ ).

Note that the use of the dot in conventional statistics is inconsistent between multiple and partial correlation. Each can be consistently expressed in  $S$ -notation, where the indicant is understood to be in sigma units, as:

	<i>S-notation</i>	<i>Conventional Notation</i>
Raw correlation	$I_{I \cdot J}$	$r_{IJ}$
Multiple correlation	$I_{I \cdot \Sigma J}$	$r_{I \cdot JKL \dots N}$
Partial correlation	$I_{I \cdot J \cdot K}$	$r_{IJ \cdot K}$
Double multiple correlation	$I_{\Sigma i \cdot \Sigma j}$	$r_{(ABC \dots D) \cdot (JKL \dots Z)}$
Part correlation	$I_{I \cdot (J \cdot K)}$	$r_{I(J \cdot K)}$ (This is not well standardized as yet.)

## S. 24, POPULATION—SCORE MATRIX (NORMALIZED)

*For analyses into components in an n-space of n variables, or, for analyses into types in an N-space of N persons*



“Correlations of people and their characteristics yield either types of persons or component characteristics”

${}^pI_I$  = Score in  $\sigma$  units of a person  $P$  (pre-superscript) in variable or characteristic  $I$  (post-subscript) (Eq. 119, Ch. VI)

$\sum_{P=1}^{P=N} {}^pI_I^2 / N$  = Variance of variable  $I$  in a population of  $N$  persons  
(Eq. 120, Ch. VI)

$\sum_{I=1}^{I=n} {}^pI_I^2 / n$  = Variance of one person's  $n$  characteristics  
(Eq. 121, Ch. VI)

$I_{I,J} = \sum_{I=1}^N {}^pI_I {}^pI_J / N$  = correlation of two characteristics, “r technic” for component analysis  
(Eq. 122, Ch. VI)

${}^{P,Q}I = \sum_{I=1}^n {}^pI_I {}^qI_I / n$  = correlation of two persons, “Q technic” for type analysis  
(Eq. 123, Ch. VI)

Instead of calculating the usual correlation between two characteristics varying among  $N$  persons, it is possible to calculate the correlation between two

persons, as they vary on  $n$  characteristics. This is called Q-technic by Stephenson (Ref. 70). In the matrix above, each row is a list distribution (see Eq. 1b, Ch. V) of one characteristic where each column is a profile of one person. Ordinarily correlation is between two rows; but in Q-technic the correlation is between two columns, showing the tendency for two person's profiles of traits to coincide, the degree to which the two persons are alike or different.

In each cell, let each axis represent the range of  $6.25\sigma$  on either side of the mean. Each person's amount of one characteristic, a cell entry, may be "normalized" either by expressing it in sigma units of deviation from the mean of the row (the usual r-technic), or in sigma units of deviation from the mean of the column (Q-technic). In the graph above both technics are simultaneously shown in plotting a vector from the cell center to a point which is determined by that person's deviation relative to other persons (an abscissa found by r-technic, Eq. 120) and his deviation relative to the rest of his own traits (an ordinate found by Q-technic, Eq. 121).

Analysis of the matrix of intercorrelations of persons by Q-technic into the analogue of components, such as principal axes or centroids, is possible. Thus the first principal axis would be the most *typical* person; the second principal axis would be the next most typical person whose characteristics were uncorrelated with those of the first typical person; the third axis would be the most typical person with the residual traits, etc. This Q-technic needs much further exploration.

## S. 25

## CORRELATION OF SCORES AND SOCIAL WORKERS' OPINIONS

<i>Total Score on Rating Scale</i>	<i>Opinions of Social Workers</i>				<i>Explanation</i>
	<i>Poor</i>	<i>Fair</i>	<i>Good</i>	<i>Excellent</i>	
25 28 31 31 33	m-21, c-4 m-26, c-2	m-21, c-10 m-23, c-8 m-25, c-8			The total score of any given home is shown at the left Its position on the horizontal scale is shown in the column heading m = material equipment score c = cultural equipment score
39 44 45 49 49 54 59 59 60 62 73 75 75		* m-30, c-15 * m-27, c-22 * m-36, c-24	m-28, c-11 m-29, c-15 m-35, c-14 m-31, c-23 m-37, c-22 m-37, c-22 m-41, c-21 m-41, c-32 m-43, c-32	* m-42, c-33	
79 80 84 85 88 98 98 99 100 107 108			* m-53, c-33 * m-61, c-36	m-54, c-25 m-53, c-27 m-57, c-27 m-52, c-33 * m-65, c-33 m-62, c-37 m-70, c-30 m-56, c-51 m-82, c-26	
TOTALS 2		6	11	10	29
OUT OF LINE— 0		* 3	* 2	* 1	6

Ref.: Chapin, F. Stuart, "A Home Rating Scale to Check Social Workers' Opinions," *Sociology and Social Research*, Vol. XIV, 1929-30, p. 13.

Descriptive formula:  $S_{25} = {}_1I_{i',j} :: {}_1I : {}_2P$

Quantic number = 0;2;0;1

Legend:

$S_{25}$  = The situation

cross-classified with

comprises

${}_1I$  = opinions of social workers expressed in 4 degrees

${}_1I$  = a total home rating indicant

and for each cell of the above

${}_1j$  = in 3 class-intervals

${}_2j$  = compared with its subratings (m and c)

${}_2P$  = the number of families is stated

### S. 26

#### CORRELATION OF THE SUM OF THE EIGHT RATINGS FOR INTELLECT WITH THE SUM OF THE EIGHT RATINGS FOR MORALITY

	14 17	18 21	22 25	26 29	30 33	34 37	38 41	42 45	46 49	50 53	54 57	58 61	62 65	66 69	70 73	74 77
10-13.....	1	1	1	1	...	1	...	...	...	...	...	...	...	...	...	...
14-17.....	2	1	...	...	1	4	...	1	...	...	...	...	...	...	...	...
18-21.....	...	...	1	...	3	2	1	...	1	1	...	...	...	...	...	...
22-25.....	...	...	...	...	2	...	1	1	1	...	1	...	...	...	...	...
26-29.....	...	...	...	...	3	5	3	...	...	1	...	...	2	...	...	...
30-33.....	...	...	3	...	2	2	3	4	2	1	2	...	...	...	...	...
34-37.....	...	...	...	1	...	4	3	4	3	1	1	2	...	1	...	...
38-41.....	...	...	...	1	1	4	3	2	1	2	3	2	1	...	...	...
42-45.....	...	1	...	1	4	2	3	5	5	3	2	3	1	...	...	...
46-49.....	...	...	...	...	...	1	5	10	6	5	2	4	3	1	...	...
50-53.....	...	...	...	...	...	...	7	9	7	4	5	1	1	1	...	...
54-57.....	...	...	...	...	...	...	2	4	17	9	5	2	4	...	...	...
58-61.....	...	...	...	...	...	1	1	4	3	10	8	3	2	1	...	1
62-65.....	...	...	...	...	...	...	...	...	3	...	3	2	3	1	...	...
66-69.....	...	...	...	...	...	...	...	...	...	...	1	3	2	1	...	1
70-73.....	...	...	...	...	...	...	...	...	...	...	...	2	1	...	...	1
74-77.....	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	...

Each entry in the table means that that number of persons had the score (sum of eight ratings) in intellect noted at the top of the column and the score in morality noted at the left of the row.

Ref.: Thorndike, Edward L., "The Relation between Intellect and Morality in Rulers," *Amer. Jour. Soc., Univ. of Chicago Press*, Vol. XLII, No. 3, Nov., 1936, p. 331.

Descriptive formula:  $S_{25} = {}_1I :: {}_1I : P$

Quantic number = 0;2;0;1

Legend:

$S_{25}$  = The situation as recorded

${}_1I$  = an indicant of morality

is composed of

and corresponding to each joint class is

${}_1I$  = an indicant of intellect

cross-classified with

$P$  = a frequency of rulers

## S. 27

A COMPARATIVE STUDY OF UNEMPLOYED AND EMPLOYED BOYS  
CLUB MEMBERSHIP COMPARED WITH UNEMPLOYMENT AND  
EMPLOYMENT

<i>Description</i>	<i>Numbers of</i>		<i>Percentages of</i>	
	<i>Un- employed</i>	<i>Em- ployed</i>	<i>Un- employed</i>	<i>Em- ployed</i>
Number who belonged to a Club. . . .	22	36	21	55
Number who did not belong to a Club . . . . .	81	29	79	45
Number who had sometime come under Club influence . . . . .	58	93	56	82
Number who had never come under Club influence . . . . .	45	20	44	18

	<i>Un- employed</i>	<i>Em- ployed</i>
Total number of boys considered in above Table equals	103	113

*Ref.:* Hughes, D. E. R., "A Comparative Study of Unemployed and Employed Boys," *Sociological Review*, Vol. XX, No. 4, Oct., 1928, p. 317.

*Descriptive formula:*  $S_{27} = {}^1_0I :: {}^1_0I_i \%P$

*Quantic number* = 0;2;0;1

*Legend:*

$S_{27}$  = The recorded situation is  
composed of

membership and club in-  
fluence

${}^1_0I$  = an all-or-none indicant of em-  
ployment

with corresponding

cross-classified with

$P$  = frequency of persons

$\%|$  = and of percentages

${}^1_0I_i$  = all-or-none indicants of club

*Comment on notation:*

This fourfold cross-classification represents the simplest type of correlation. It measures the degree of truth in such a common sociological statement as, "Club members tend to be employed."

## S. 28

## 2. A UNIVERSE OF 108 ELEMENTS

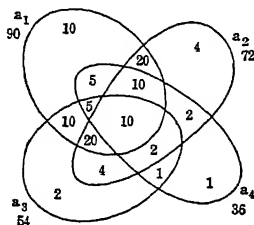
As a first example, let

$$A = 108; \quad a_1 = 90; \quad a_2 = 72; \quad a_3 = 54; \quad a_4 = 36.$$

Then very many kinds of possible linkages between the four variables are possible, and numerous different tetrads of correlation coefficients. But if  $a_1$  and  $a_2$  are really random samples, that is, independently formed samples, of  $A$ , then  $a_2$  will have *most probably*

$$\frac{72}{108} \times 90 = 60$$

elements in common with  $a_1$ . Similarly  $a_3$ , which has 54/108ths of the whole  $A$ , will have *most probably* that same fraction of each of the categories formed by the existing variables  $a_1$  and  $a_2$ ; and so on. In short, the most probable distribution of linkages is as follows:



(The total number within the diagram is 106, *i.e.* 2 elements, *most probably*, escape being involved in any one of the four variables.)

The six correlation coefficients which this most probable structure would give are

	$a_1$	$a_2$	$a_3$	$a_4$
$a_1$		.75	.65	.53
$a_2$	.75		.58	.47
$a_3$	.65	.58		.41
$a_4$	.53	.47	.41	

wherein each  $r$  is given by

$$\frac{\text{common elements}}{\text{geometrical mean of two totals}},$$

$$i.e. \quad r_{12} = \frac{60}{\sqrt{90 \times 72}} = 0.75.$$

*Descriptive formula:*  $S_{23} = ( (\Sigma I)_i \bullet (\Sigma I)_{j,i} )_{i,i}$

*Quantic number* = 0;2;0;0

*Legend:*

$S_{23}$  = The situation  
is composed of

$(\Sigma I)_{j,i}$  = another similar index (not  
itself)

$()_i$  = an index

and repeated for each of

which is

$|_i$  = 4 indices

$\Sigma I$  = a sum of components

$|_{i,i}$  = cross-classified with each  
other

and is

$\bullet$  = correlated with

*Comment:*

The situation is a general statistical one and is a sociological situation only if it represents characteristics of people. The formula at the bottom (see Eqs. 13a and 43, Ch. VI) is accurate only for an infinite population. For a small population it will be only approximate within the limits of the standard error of sampling.

The components, represented by overlapping areas, which add up to yield the four intercorrelated indicatory indices here, are all synthesized from 108 independent equal elements in the "most probable" pattern. This pattern gives an intercorrelation matrix of rank one, which can be reanalyzed into one general component,  $g$ , common to the four indices and four specific components. These five components are transformations of the components, diagramed above as the common elements in overlapping areas of the circles.

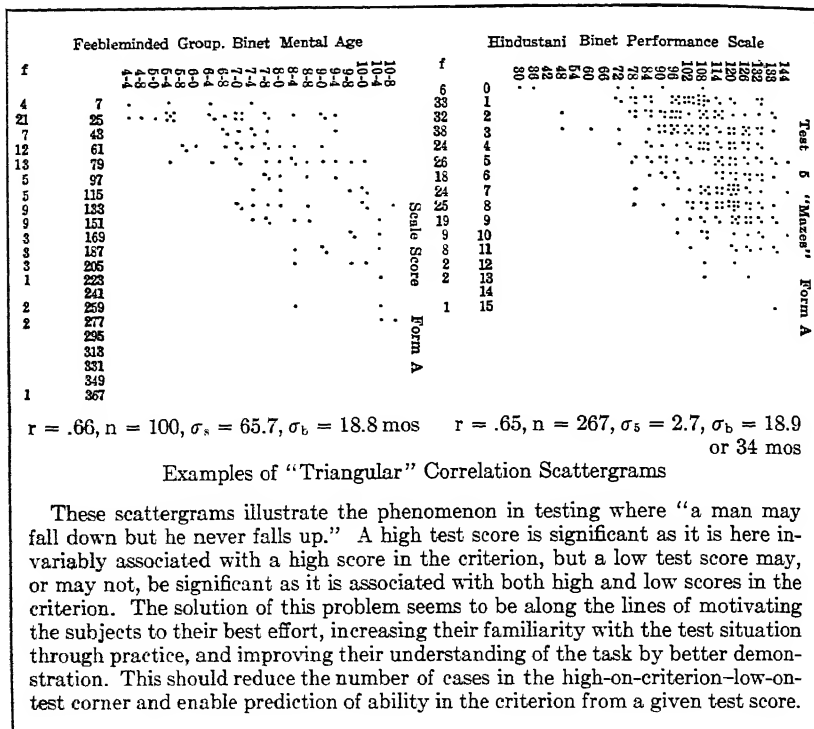
*Comment on notation:*

1. Without specifying the composition of the indices, the matrix of intercorrelations alone could be symbolized by the descriptive formula,  $S = I_{i,i}$ , denoting the correlation coefficients of  $|_i$  indicants with each other.

2. Since a correlation is a scalar product its quantic formula involves an exponent of 2. It can readily be shown that the common elements formulae for correlation (see S. 7, Ch. VI) are but simplifications of covariances, as in Eqs. 14-19, Ch. VI, which, vectorially, are scalar products.

Whether the second cross-classifying in the matrix of intercorrelations should step the exponent up in the quantic formula is not clear. This awaits further research. For simplicity and also because the matrix does not appear to be summarizable in any form equivalent to a scalar product, the exponent in the quantic formula is for the present left as 2.

## S. 29



Ref.: Dodd, Stuart C., *International Group Mental Tests*, Princeton University Store, 1926, p. 70.

Descriptive formula:  $S_{29} = ({}^e I :: {}^s I : P)_P$

Quantic number = 0;2;0;1

Legend:

$S_{29}$  = The situation

$::$  = cross-classified

comprises

in each of

$I, I$  = indicants of intelligence

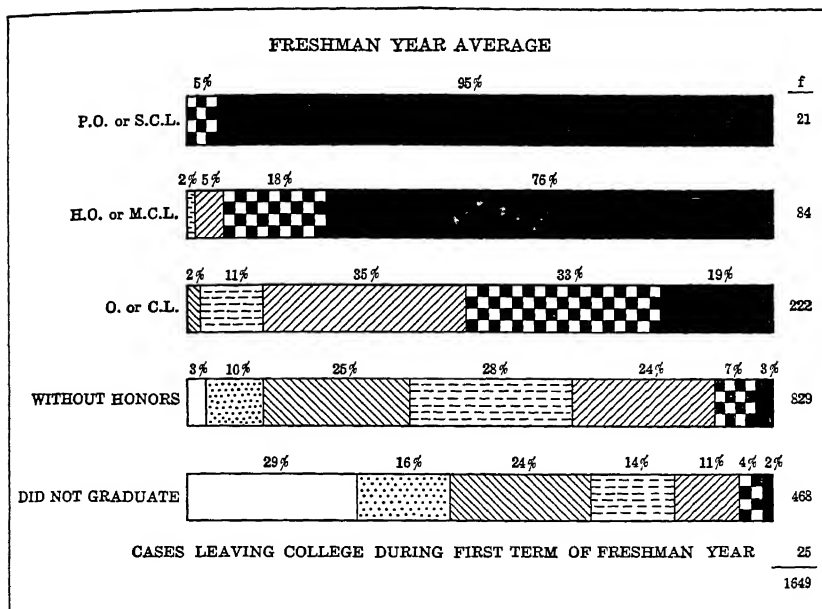
$P_P = 2$  popula-  
tions

$i, i$  = subdivided into their class-  
intervals

{ a feeble-minded  
plurel  
a Hindu plurel

${}^e$  | = (and with their sigma ranges  
also stated)

## S. 30



Ref.: Department of Personal Study, *School and Society*, Yale Univ., July 18, 1934.

Descriptive formula:  $S_{30} = ({}^1I :: {}_iI : \%P)$

Quantic number = 0;2;0;1

Legend:

$S_{30}$  = The situation

$|$  = the class-intervals (shadings)

is a record of

of

${}^1I$  = 5 ordinal points of the graduation-status indicant

$I$  = Freshman year standing with the corresponding

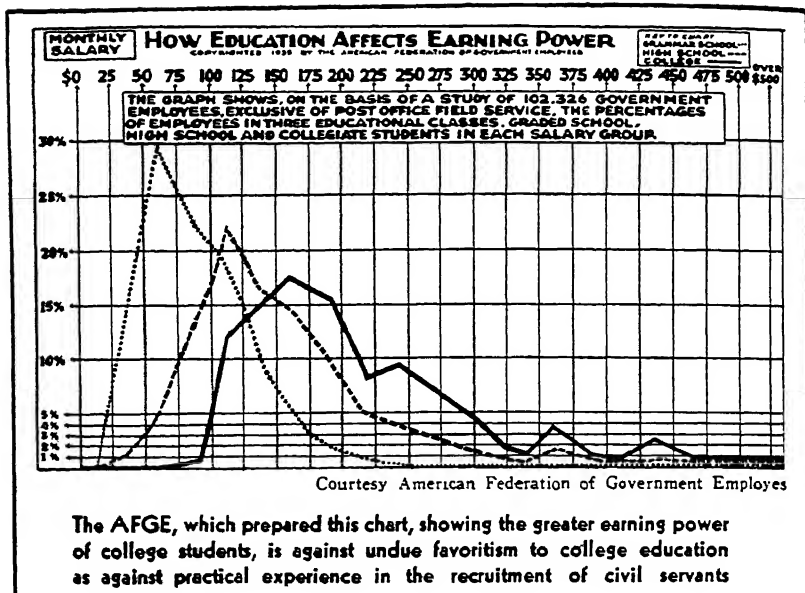
$::$  = cross-classified with each of

$\%P$  = percentage frequency of Yale students

Comment:

When the horizontal bars tend to be each of one kind of shading the correlation nears perfection; when each shading tends to run in a vertical band through all the horizontal bars, the correlation is near zero. As the step-like diagonal boundaries between any two shadings go from horizontal to vertical the correlation drops.

## S. 31



Ref.: Weybright, Victor, "Our Civil Servants," *Survey Graphic*, Vol. XXV, No. 2, Feb., 1936, p. 95.

Descriptive formula:  $S_{31} = {}_1I :: ({}_jI : {}_pP)$

Quantic number = 0;2;0;1

Legend:

$S_{31}$  = The situation is a  
record of

(:) = distributions of  
 ${}_jI$  = salaries in 20 \$25 classes  
with corresponding

${}_jI$  = 3 class-intervals  
of

${}_pP$  = percentage frequencies of gov-  
ernment employees

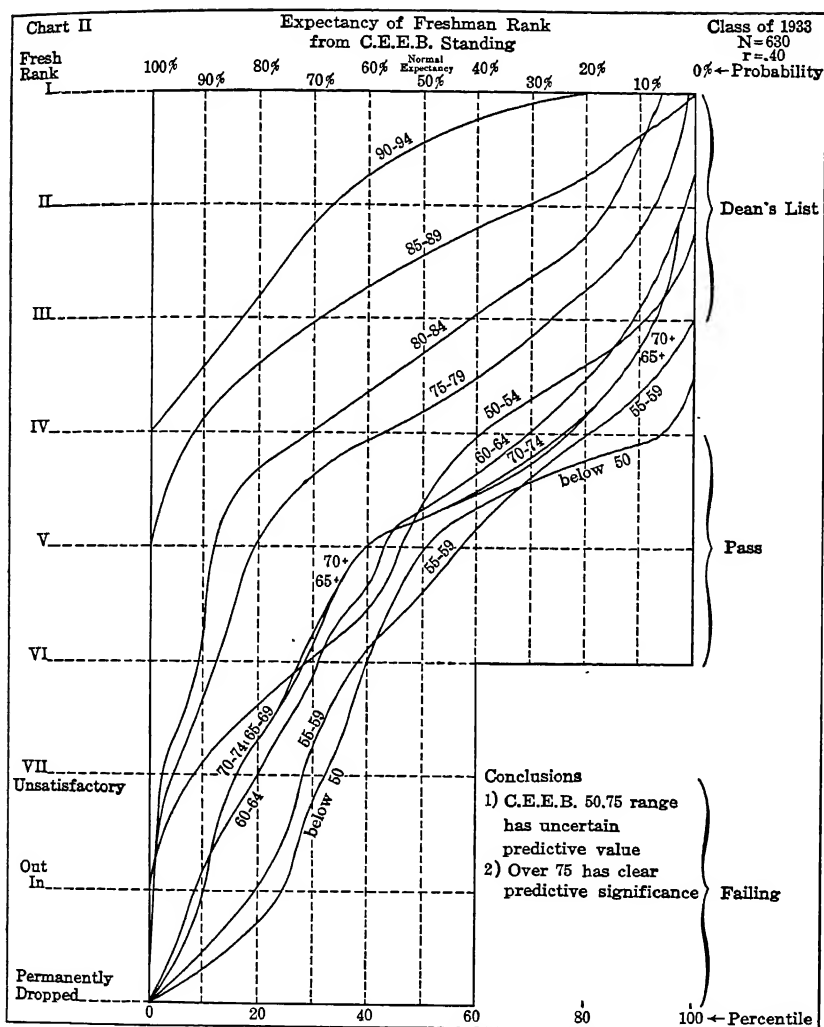
$I$  = schooling

:: = cross-classified with which are

Comment on notation:

The fact that the graph is a correlation (between indicants of earnings and of schooling) is obscured by the graph's being presented in the form of frequency distributions. These, however, are but three slices of the correlation solid of S. 5 as presented in S. 4, Ch. VI.

## S. 32



Ref.: Dodd, Stuart C., *A Report to the Deans of Harvard College* (unpublished), 1935.

Descriptive formula:  $S_{32} = {}_1I : {}_1I : {}_{2\%}P$

Legend:

Quantic number = 0;2;0;1

$S_{32}$  = The situation

is a record of

trance Examination standing  
("C.E.E.B.")

${}_1I$  = 9 class-intervals of Freshman rank

$:$  = and, for each joint reading,

${}_{2\%}|$  = the cumulative percent

$::$  = cross-classified

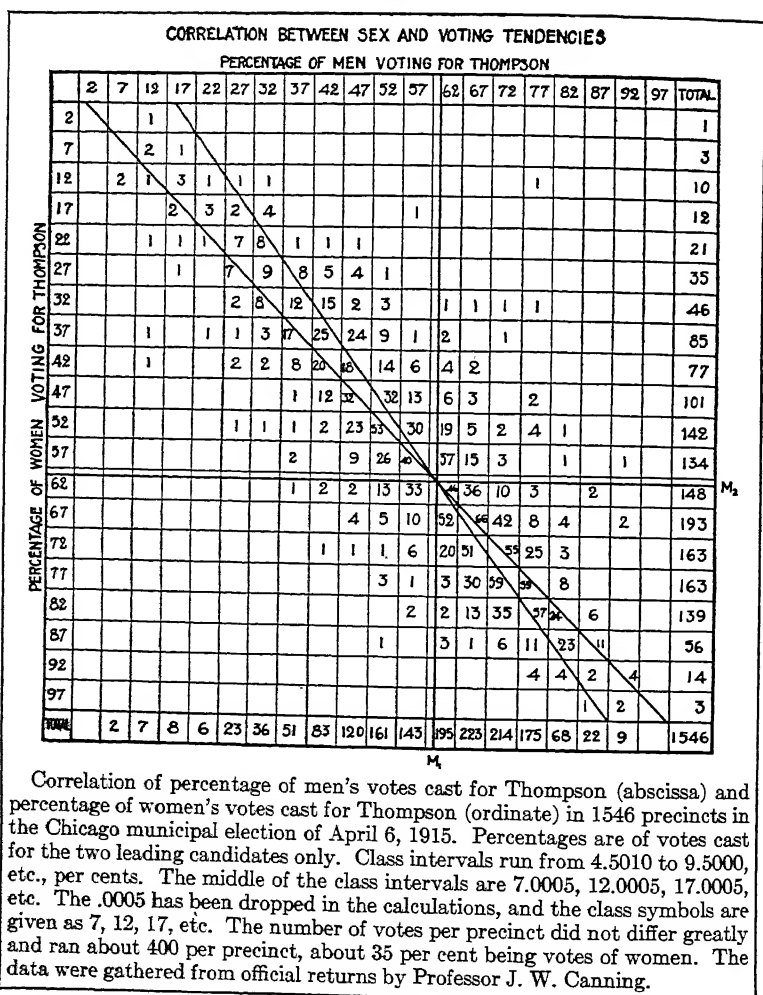
of

with

$P$  = the class of 1933

${}_1I$  = 9 class-intervals of College En-

## S. 33



Ref.: Kelley, Truman L. *Statistical Methods*, Macmillan Company, 1923, p. 183, Chart XXI.

Descriptive formula:  $S_{33} = {}_1(I) :: {}_1(I) : {}_2P$

Legend:

$S_{33}$  = The situation

records

${}_1(I) = ({}_2P_r) =$  an index of male voting (the % of men voting for Thompson)

cross-classified with

${}_1(I) = {}_2P_{r'}$  = an index of female

Quantic number = 0;2;0;1

voting (the % of women voting for Thompson)

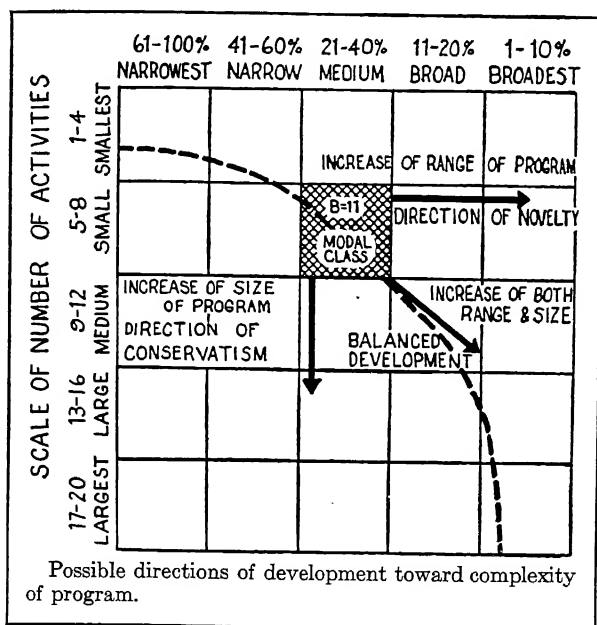
with

${}_2P$  = the frequency of precincts at each joint class-interval of the indices

*Comment on notation:*

The population ratios are written as indices because the point of the situation is definitely the characteristic definable as "voting of the sexes," and not the size of any plurel. Since the graph is a correlation and not an interrelation (see Ch. VII) the cross-classification of indices is a more accurate description than the cross-classification of people. A population percentage is also a mean of an all-or-none indicant.

## S. 34



Ref.: Chapin, F. Stuart, *Contemporary American Institutions*, Harper and Brothers, 1935, p. 189.

Descriptive formula:  $S_{34} = {}_iI :: {}_j(|_{z_1}) : {}_pP$

Legend:

$S_{34}$  = The situation

is composed of

${}_iI$  = 5 class-intervals of an indicant of the number of church activities

$::$  = cross-classified

with

$(|_{z_1})$  = the number of kinds of activities

${}_j|$  = expressed in 5 class-intervals

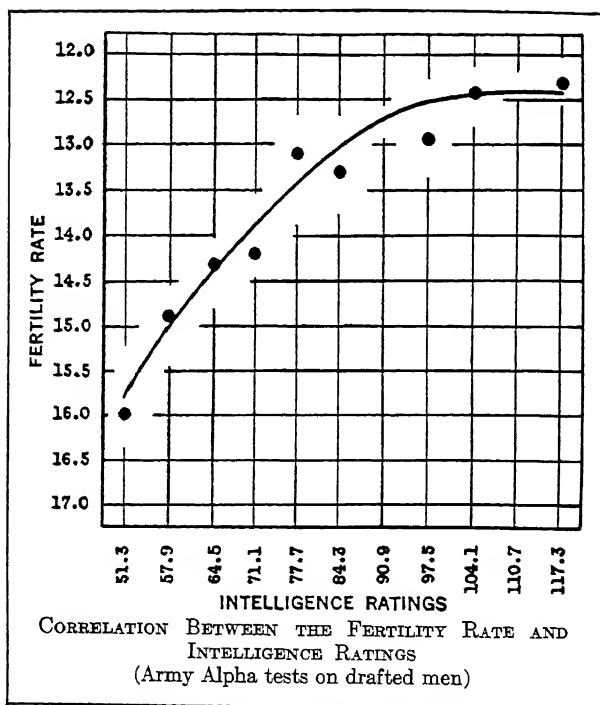
$:$  = with the corresponding

${}_pP$  = number of churches

*Comment:*

The index  $(|_{z_1})$  measures the "differentiation" of any church's activities, as defined in Chapter V.

## S. 35



Ref.: Ogburn and Tibbits, "The Birthplace and the Social Classes," *Social Forces*, Vol. III, Sept., 1929, p. 8.

Descriptive formula:  $S_{35} = 'x'I : \Sigma(I)P^{-1}$

Quantic number = 0;11;0;9

Legend:

$S_{35}$  = The situation

records

I = an indicant of intelligence

in

$|$  = 11 class-intervals

'x' = between the limits of 51  
and 117 in army Alpha  
scores

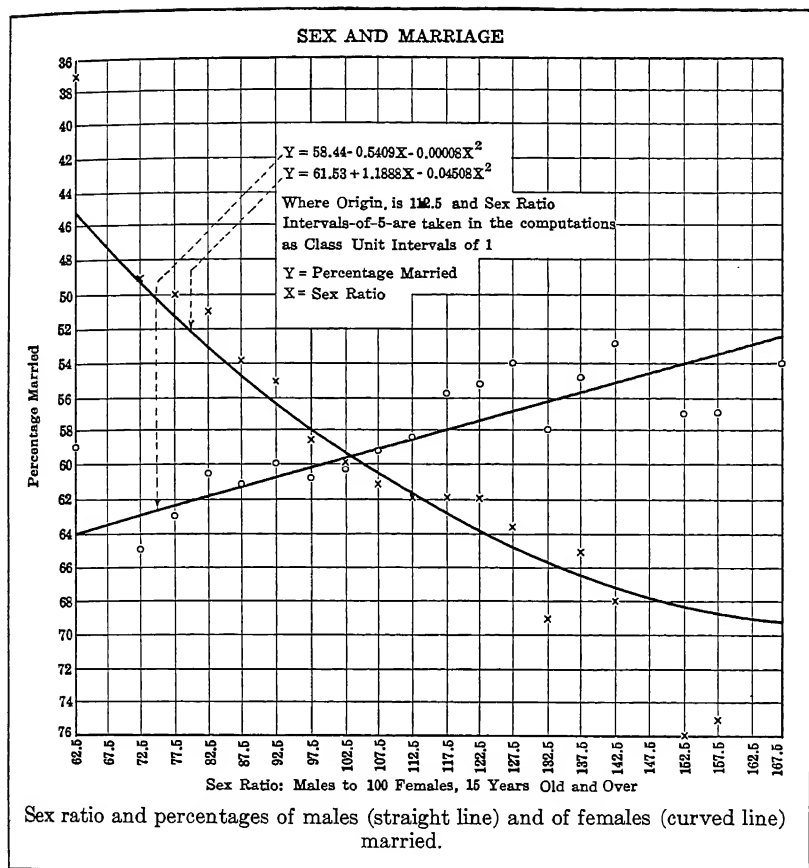
: = each with corresponding

$\Sigma(I)P^{-1}$  = mean index of fertility

Comment on notation:

Two alternative analyses present themselves. The index of fertility (I) is a population rate,  $\%P$ , though its factors are not specified. Since the point of the graph is a correlation of the two characteristics, intelligence and fertility, the latter is written as an index, denoting a compound characteristic. The mean index of fertility for each intelligence array might be written  $\bar{M}(I)$ , reducing the quantic to I·I and the quantic number to 0;2;0;0. But following the notational rule that, if securing the data required counting persons, then the P should be written explicitly, the descriptive formula has been written as above. The situation fundamentally is a correlation of two characteristics of a population (in the form of a regression curve of one on the other), and the quantic classifies it as such.

## S. 36



Ref.: Groves and Ogburn, *American Marriage and Family Relationships*, 1928, Henry Holt, p. 145.

Descriptive formula:  $S_{36} = \underline{P}_D : {}^{-1}\tau_i(I) : (I)_{,,}$

Quantic number = 0;11;0;1

Legend:

$S_{36}$  = The situation

the limits of 60 and 170  
with the corresponding

records for each of

$\underline{P}_D$  = the 2 sex plurals

$(I)_{,,}$  = index of marriage (= 100  
 $\Sigma^{1.0} I/P$  = percent married =  
the mean of an all-or-none  
inditant)

$(I)$  = the sex ratio (=  $P/P_{,,}^{-1}$  =  
males per 100 females)

${}^{-1}\tau_i$  = in 22 class-intervals between

Comment:

As in the previous graphs, the population ratios are treated as indices of characteristics, since the point of the graph is the correlation of these characteristics of a population, and not the size of the population or its interrelations (as defined in Ch. VII).

## S. 37

## RESEMBLANCE OF SIBLINGS IN MENTAL TRAITS

	Brothers	Sisters	Brothers and Sisters
Veracity . . . . .	.47	.43	.49
Assertiveness . . . . .	.53	.44	.52
Introspection . . . . .	.59	.47	.63
Popularity . . . . .	.50	.57	.49
Conscientiousness . . . . .	.59	.64	.63
Temper . . . . .	.51	.49	.51
Ability . . . . .	.40	.47	.44
Handwriting . . . . .	.53	.56	.48
Mean . . . . .	.52	.51	.52

Ref.: Holmes, Samuel J., *The Trend of the Race*, Harcourt Brace and Co., 1921, p. 104.

Descriptive formula:  $S_{37} = {}^m_1({}_pP^{-1} \Sigma {}^{p'}_q I^{q'} I)_{i;q}$

Quantic number = 0;2;0;9

Legend:

$S_{37}$  = The situation

$|_i$  = 8 characteristics

is a record of

in

( ) = correlation coefficients

${}_p I$  = for one characteristic (in  
sigma units)

$|_q$  = 3 kinds of { brothers  
plurels { sisters  
siblings

${}^{p'}|, {}^{q'}$  = between the two persons

and

of

${}^m_1$  = mean coefficients are also  
given

${}_p P^{-1}$  = sibling pairs

calculated for

Comment on notation:

The correlation coefficient being an index in S-notation is enclosed in parenthesis. With less specification of detail it could be equally well written  $S = {}^m_1(I)_{i;q}$ , or  ${}^m_1 r_{i;q}$ , with the legend stating the index to be a correlation coefficient.

## S. 38

MULTIPLE AND PARTIAL CORRELATION RATIOS BETWEEN THE TOTAL STANDARD OF LIVING (TOTAL EXPENDITURES PLUS VALUE GOODS FURNISHED) ( $X_1$ ), AND AGE OF OPERATORS AND WIVES ( $X_2$ ), NET CASH INCOME ( $X_3$ ), MILES FROM VILLAGE ( $X_4$ ), AND EDUCATION OF OPERATORS AND WIVES ( $X_5$ )

<i>Subscripts</i>	<i>eta</i>	<i>eta squared</i>	<i>Subscripts</i>	<i>eta</i>	<i>eta squared</i>
1.2345	.53	.28	13.245	.46	.21
1.345	.53	.28	14.235	.32	.10
12.345	0	0	15.234	.28	.08

*Ref.*: McCormick, Thomas C., "Owner-Tenant Contrasts," *Sociology and Social Research*, Vol. XVII, No. 2, Nov., 1932, p. 157.

*Descriptive formula*:  $S_{33} = (I)^2 \cdot \Sigma_j$

*Quantic number* = 0;2;0;9

*Legend*:

$S_{33}$  = The situation

stating

presents

$\cdot$  = the correlation between

(I) = indices

$|_j$  = one indicant

$|^2$  = and their squares

$\Sigma_j$  = and the sum of four others

*Comment on notation*:

The dot in S-notation denotes correlation between the two indices it separates. This usage is the same as in conventional statistics in denoting the multiple correlation and in vector algebra, where the dot denotes the scalar or dot product (equivalent to correlation in statistics). Partial correlation may be distinguished in S-notation by the subtraction symbol, as in  $I_{j,-i}$ , meaning the correlation of I, and  $I_j$ , after eliminating j other indices,  $I_j$ . "Part correlation" can be readily distinguished from partial correlation by a parenthesis, thus  $I_{j,(-i)}$ , meaning the correlation of the index I, with the index  $I_j$ , after the latter index has had the effect of the other indices,  $I_i$ , eliminated.

Since the situation presents correlation of characteristics in a population the quantic number may be written 0;2;0;9, even though the P is not explicitly written, as it is required in calculating the index. The formulae for eta involve P as well as the squares of the indicants in obtaining the sigmas by arrays. (See S. 3, Ch. VI.)

## S. 39

INTERCORRELATION OF THE SECTION SCORES IN THE RURAL  
AND RURAL RANGES

<i>Section</i>	II	III	IV	V	VI	<i>Scale</i>
II. Remedies	<b>.82</b>	.05	.32	.41	.11	.66
III. Infants	.77	<b>.82</b>	.07	.15	-.20	.42
IV. Food and Cleanliness	.85	.85	<b>.76</b>	.34	.10	.68
V. Insects	.84	.70	.82	<b>.84</b>	-.14	.67
VI. Housing	.87	.85	.76	.84	<b>.96</b>	.23
Scale	.92	.93	.93	.94	.97	<b>.92</b>
Average intercorrelation	.83	.79	.82	.80	.83	

The coefficients in the upper right triangle of the table are from the rural range (b + c sample N = 100. Scale S.D. = 43) while those from the lower left triangle are from the rural range (b + c + U sample, N = 150, S.D. scale = 175). The coefficients in heavy type in the main diagonal cells are reliability correlations between male and female informants (rural range, S.D. scale = 120, "r" sample, N = 45).

*Ref.: Dodd, Stuart Carter, A Controlled Experiment on Rural Hygiene in Syria, American Press, Beirut, 1934, Table 5, p. 79.*

*Descriptive formula:*  $S_{39} = {}^m(I)_{p:i}$

*Quantic number* = 0;2;0;1

*Legend:*

$S_{39}$  = The situation  
records

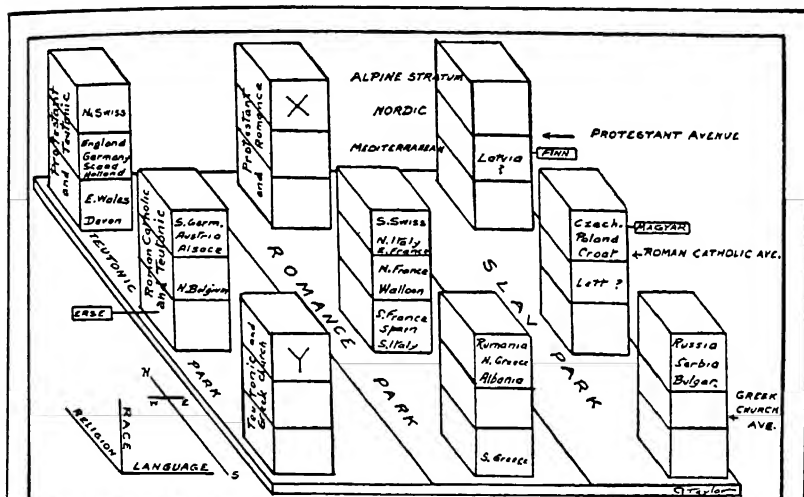
$|_{i:i}$  = the intercorrelations  
of

$|_p$  = for each of  
3 plurels

{ rural  
"rural"  
"male and  
female"

(I) = 6 indicants of hygiene  
 ${}^m|$  = and average intercorrelations  
are also stated

## S. 40



Three-dimension diagram showing the combinations of the three variables—race, language, and religion—which characterize the nations of Europe. In each "building" the lowest stratum is occupied by "Mediterranean" stocks, the middle stratum by Nordic, the top stratum by Alpine stocks. For example: the Northern Swiss being Alpine (top stratum), Protestant (north avenue), and Teutonic speakers (west park), are found in the top of the left-hand "building." X (Protestant-Romance) and Y (Teutonic-Greek Church) are virtually absent among European peoples.

Ref.: Taylor, Griffith, "Environment and Nature," *Amer. Jour. Soc.*, Vol. XL, No. 1, July, 1934, p. 34.

Descriptive formula:  $S_{40} = I^0 :: I^0_j :: I^0_k : P_p$

Legend:

$S_{40}$  = The situation  
records

$I^0_j$  = 3 languages

$::$  = cross-classified by

$I^0_j$  = 3 races

$::$  = cross-classified by

$I^0_k$  = 3 religions

$:$  = with corresponding

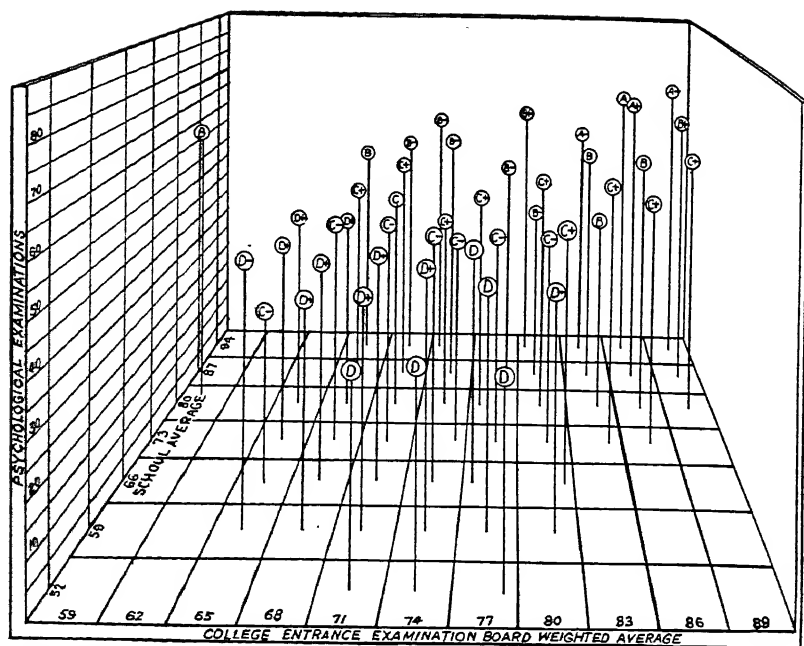
$P_p$  = national plurels in Europe

Quantic number = 0;0;0;1

Comment:

This situation has a quantic number of 0;0;0;1 which classifies it under Plurels (Ch. IV). It is presented with the situations of correlation in this chapter, however, to illustrate the effect of the operation of observing, including computing, in determining the quantic. If the data were more penetratingly and fully observed to the level of computing contingencies, the qualitative characteristics denoted by  $I^0$  would be thereby converted to correlated characteristics denoted by  $I^2$  (or  $I^3$  for the present case of a triple correlation).

## S. 41



Ref.: *Scholastic Aptitude Tests, A Manual for the Use of College Officers, prepared by the College Entrance Examination Board, p. 10.*

*Descriptive formula:*  $S_{11} = {}_i I :: {}_j I : (\Sigma I P^{-1})_k$

*Quantic number* = 0;3;0;9

*Legend:*

$S_{11}$  = The situation

$\Sigma I P^{-1}$  = means of students

records

on

${}_i I$  = the College Entrance Examination indicant

$::$  = cross-classified with

${}_j I$  = a High School standing indicant

$:$  = with corresponding

$|_k = 2$  indices

{ Psych. exam.  
(height of  
lines)  
College marks  
(letters in  
circles)

*Comment:*

This situation illustrates a triple correlation, the cross-classification in a three-dimensional space of three characteristics. The swarm of circles-on-pins represent the ellipsoidal scatter diagrammed in S. 23, Ch. VI.



S. 42 (*Continued*)

*Descriptive formula:*  $S_{42A} = (\sigma I)_{i \cdot i}$

*Legend:*

$S_{42A}$  = The matrix A  
is

A. *Quantic number* = 0;2;0;0

$(\sigma I)_{i \cdot i}$  = the correlation coefficients (I)  
 $|_{i \cdot i}$  = of all 4 tests with each other

*Descriptive formula:*  $S_{42B} = (I = JK_{\Sigma i})_i$

*Legend:*

$S_{42B}$  = The situation B  
records

$I_i$  = 4 tests (" $x_1, x_2, x_3, x_4$ ")  
each equal to

B. *Quantic number* = 0;1;0;0

$J_{\Sigma i}$  = the sum of 4 components (y's)  
K = each times a numerical coefficient

*Descriptive formula:*  $S_{42C} = (J = KI_{\Sigma i})_i$

*Legend:*

$S_{42C}$  = The situation C  
records

$J_j$  = the four components (y's)  
each expressed as

$I_{\Sigma i}$  = a sum of 4 test scores  
K = each times its coefficient

C. *Quantic number* = 0;1;0;0

*Descriptive formula:*  $S_{42D} = \Sigma (I)_{i \cdot j}$

*Legend:*

$S_{42D}$  = The situation D  
records

$(I)_{i \cdot j}$  = the variances (v) and co-  
variances (p)  
of

$|_{i \cdot j}$  = the 4 components with each  
other

$\Sigma$  = and their sum is also stated

D. *Quantic number* = 0;2;0;0

*Descriptive formula:*  $S_{42E} = |(I)_{i \cdot j}| = |(I)_{i \cdot i}| = .2353$

*Legend:*

$S_{42E}$  = The situation E  
is an equation between

$|(I)_{i \cdot j}|$  = the final determinant of the  
intercorrelations of the 4  
components (y)

and  
 $|(I)_{i \cdot i}|$  = the initial determinant of  
the intercorrelations of the  
4 tests (x)

E. *Quantic number* = 0;2;0;0

F. The right hand half of S. 42 presents situations of identical form but with differing quantities.

*Comment:*

The two pages quoted from Kelley compare the analysis of a set of 4 indices ("Reading Speed," "Reading Power," "Arithmetic Speed," "Arithmetic Power," in a 7th grade population of 140) into components, first, by the

Hotelling-Kelley technic into principal axes, and second, by Thurstone's technic into centroids. The comparison indicates that the latter technic has distorted the data, since the final determinant does not exactly equal the initial determinant. The two technics have yielded variances of the components,  $v(y)$ , which are closely alike. As may be seen from the final determinants, the components are independent of each other as they have zero intercorrelations; and they are arranged in order of descending contribution to the total variance of the situation (1.84, 1.46, .52, .17).

## S. 43

If  $n$  is the number of variables there will be  $n - 2$  boundary conditions. The first boundary condition will show how far the set of correlations can be raised as a whole without the use of any group factor common to more than two of the variables. The second boundary condition will show how much further they can be raised without the use of any group factor common to more than three variables, and so on for the third, fourth, and all the remaining boundary conditions until we come to the last, which shows how far the correlations can be raised without the use of a general factor. All these conditions can be conveniently expressed in the form of a determinant as follows:—

$$B_{n,k} = \begin{vmatrix} k & r_{12} & r_{13} & \cdots & r_{1n} \\ r_{21} & k & r_{23} & \cdots & r_{2n} \\ r_{31} & r_{32} & k & \cdots & r_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & r_{n3} & \cdots & k \end{vmatrix} = 0 \dots \dots \dots (1)$$

in which  $k$  takes the values  $-1, -2, -3$ , etc., respectively, according as the first, second, third, etc., boundary condition is required, and the  $r$ 's are the (positive) correlations. For the sake of uniformity of interpretation with regard to signs it is advisable to regard the term which is the product of all the  $k$ 's as positive—that is to say, reverse the signs when  $n$  is odd. The boundary conditions can then all be read; if  $B$  is positive the higher order factors are not necessary. The symbol  $B$  replaces the  $D$  of previous papers, so as to include the term involving the product of all the  $k$ 's. The general expression was reached by the writer in expanded form by the method described in the Appendix: for its subsequent expression in the form adopted in equation (1) he is indebted to Mr. T. B. Black, of Edinburgh University.

Descriptive formula:  $S_{43} = B_{n,k} = |(I)_{i,i}| = 0$   
 $(I)_{i,i} = -1, -2, -3,$   
 $\dots, -n$

Quantic number = 0;2;0;0

Legend:

$S_{43}$  = The situation  
 comprises

$|_{i,i}$  = of  $i$  ( $= n$ ) indicants with  
 each other

$|| = 0$  the discriminant  
 of

when

$(I)$  = intercorrelations

$(I)_{i,i}$  = the main diagonal cells con-  
 tain the successive negative  
 integers

Comment on notation:

The vanishing of this discriminant (within standard error limits) is the critical test or boundary condition as to whether components of higher or lower order are compulsorily present. For a simpler, but less accurate, boundary condition see Eqs. 109 and 110 in connection with S. 21, Ch. VI.

#### IV. NOTES

1. Chapter III on Indicators dealt with the stratum where there was no direct reference to population (the  $P^0$  stratum) and in that stratum with the static and spaceless array,  $T^0 L^0$ . It dealt with indicators to various powers, chiefly the attribute array,  $I^0$ , of qualitative characteristics, and the indicant array,  $I^1$ , of quantitative characteristics. The array over that in the next higher stratum containing the population sector,  $P^1$ , was then taken up in Part III. The first chapter (Chapter IV) in Part III took up the combination of attributes and population—forming Plurels,  $I^0 P^1$ . The following Chapter V took up the combination of indicants and population which defines Distributions,  $I^1 P^1$ . In our accumulative approach we now take up the combination of population and more than one indicant which defines Correlation,  $I^2 P^1$ .

2. The general descriptive formula for a correlation scatter would thus be

$$(I) :: {}_j(I) : (I)_k \quad (\text{Eq. 2, Ch. VI})$$

where the indices  $(I)$  (as defined in Chapter III) denote any sum or product of indices from any sector.

3. In a more general formula which includes curvilinear correlation

$$\eta_{IJ} = \sigma(\sum_j I_j) / \sigma I = \text{sigma of means of } I_j \text{ arrays for each of the } {}_j J \text{ class-intervals}$$

in terms of the sigma of  $I$ , the correlation ratio of  $I$  on  $J$

(Eq. 4, Ch. VI)

(and in its paired ratio  $\eta_{JI}$ ) the indicant is squared in computing the sigma of means (which is then expressed in terms of the sigma of the distribution to bring  $\eta$  to a zero-to-unity scale). Therefore,  $I^2$  is the proper quantic, i.e., power of the indicator,  $I$ , here also. (The correlation ratio can also be stated in an alternative form in terms of the mean sigma of arrays. Study S. 3, Ch. VI.)

4. As explored in more detail in the twenty-four diagrams interpreting cor-

relation that are appended to this chapter, there are many variant forms of correlation formulae. Some are variants suited to the form of the data (as in biserial and tetrachoric correlation). Others are variant methods of computing (such as the product vs. the difference formulae). All these can be shown to involve the essential principle, however, of the cross-product moment or its equivalent. The less accurate first-order correlation formulae like  $Sm$ , the coefficient of Similarity (Ref. 57, p. 659), and Spearman's footrule from gains in rank are not included in the preceding statement which covers only the more accurate second-order correlation.

5. The more detailed formulae are:

For any scalar product:

$$\bar{I} \cdot \bar{J} = IJ \cos \theta \quad (\text{Eq. 6a, Ch. VI})$$

for a scalar product of unit vectors:

$$\bar{I} \cdot \bar{J} = \cos \theta \quad (\text{Eq. 6b, Ch. VI})$$

Since the scalar product is computable as the covariance (i.e., the average cross-product):

$$\bar{I} \cdot \bar{J} = \Sigma IJ/P = \sigma_I \sigma_J r_{IJ} \quad (\text{Eq. 5b, Ch. VI})$$

which in sigma units becomes

$$\bar{I} \cdot \bar{J} = \Sigma IJ/P \sigma_I \sigma_J = r_{IJ} \quad (\text{Eq. 5c, Ch. VI})$$

Since vectors can be handled as algebraic quantities, the overlining denoting the geometric interpretation can be omitted as in Eq. 5.

Note that in the limiting case where the population is one person, ( $P = 1$ ), the mean product as stated in Eq. 3, Ch. VI becomes the ordinary arithmetic product  $I \cdot J$ . Thus the arithmetic product is the special case at the limit of the scalar product. The scalar product or correlation coefficient is both more general and far more useful for sociological purposes.

6. Whether a probability is simply a ratio of frequencies or whether it involves much more than that, such as "a degree of rational belief," is a metaphysical question dividing philosophical statisticians into schools. (See Ref. 64, p. 281, footnote and further references.) The ratio, which is the objective measure of probability, is the fact at the scientific level, for science requires objectivity and agreement among observers. The discussion at the more metaphysical level of interpretation, however, should be known to the scientist for more complete understanding of the phenomena he observes.

7. When the cases are persons this may be simply written as an attribute population product (see Ch. IV), as:

$${}_P P = \text{a population probability} \quad (\text{Eq. 7b, Ch. VI})$$

where  $P$ , are the persons with the implicit quality,  $I^0$ , and  ${}_P P$  is the percent of persons with that quality. (See the graphed situations in Chapter IV for examples.) Similarly when the cases are occasions or temporal dates when some event did or did not happen, the attribute-time products give the ratio as:

$${}_t T^0 / ({}_t T^0 + {}_u T^0) = \frac{t}{t+u} = \text{a temporal probability} \quad (\text{Eq. 7c, Ch. VI})$$

This is condensed notation from  ${}^1I^0 : T^0 = {}^uT^0$  and  ${}^1I^\infty : T^0 = {}^uT^0$  which defines  ${}^1$  as the number of dates characterized by the presence of the quality and  ${}^u$  as the number of dates characterized by its absence.

When the attribute, instead of being all-or-none, is subclassified  $I_i^0$ , the probability of each class is the ratio of the number of cases in it to the whole number of cases. Letting an indicant represent the number of cases or frequency of an attribute so that

$$\Sigma {}^1I^0 = I^{+1} \text{ and } \Sigma {}^1I_{i';(j',j'',j''',\dots)}^0 = I_{i';(j',j'',j''',\dots)} \quad (\text{Eq. 8a, Ch. VI})$$

then  $I_{j'/j'} + I_{j''/j''} + I_{j'''/j'''} + \dots = 1.00$  probabilities of parts of a whole  
(Eq. 8b, Ch. VI)

Thus, in S. 11, Ch. II showing a university enrolment by nationalities, the probability of any student met at random being an Abyssinian is

$$P_{q'/p'} (= I_{j'/j'}^0 P_0) = 2/1421;$$

or of being a Palestinian is

$$P_{q''/p'} (= I_{j''/j''}^0 P_0) = 238/1421$$

The probability of *any* one of several classes is the sum of their separate probabilities

$$I_{(j'+j'')/j'} = \text{the probability of either } j' \text{ or } j'' \quad (\text{Eq. 9, Ch. VI})$$

The probability of some case in any one of all of the classes is complete certainty, of course, as expressed by the unity in Eq. 8b.

8. It may be noted that a lower case letter script denotes an aggregation (i.e., a list of entities) only when in the scriptal position. It denotes a single number of entities when written as a base letter (as in the right-hand member of Eq. 7a, Ch. VI), or when, in the scriptal position, it is preceded by the summation sign,  $\Sigma$ , which directs the operator to sum the entities and convert the list to a single number as in Eq. 7c, Ch. VI above.

9. The binomial expansion is given by:

$$(p + q)^n = \sum_{r=0}^{n+1} \frac{n!}{(n-r)!r!} p^{n-r} q^r = \text{the binomial expansion, } r = 0, 1, 2, 3, \dots n \quad (\text{Eq. 10, Ch. VI})$$

$$\text{where } p + q = 1 \left( \text{i.e., where } p = \frac{i}{i+j} \text{ and } q = \frac{j}{i+j} \right), \quad (\text{Eq. 10a, Ch. VI})$$

and where the observed fact is the sum of the all-or-none elements,  $n$  in number. When  $n$  is large the resulting binomial distribution approaches the normal distribution curve. If  $p = q$  (i.e., if  $i = j$ ), the curve is symmetric and it is skewed in proportion to their inequality.

$$\frac{n!}{(n-r)!r!} = {}_nC_r = \text{the number of combinations of } n \text{ things taken } r \text{ at a time.} \quad (\text{Eq. 10b, Ch. VI})$$

10. This reversal of reasoning is logically possible but not necessarily true, as it passes from the proposition, "If elements exist, then a normal curve" to the converse proposition, "If a normal curve, then elements exist." Therefore, the existence of elements inferred from the given fact of a normal curve is an hypothesis.

11. This assumption of a very large number of elements was used by Gauss in deriving the normal curve, which is often called the Gaussian curve after its discoverer. In the derivation of probable error formulae, differentials are taken as the small elements and squared and averaged and assumed to be normally distributed. In deriving the first formula for the normal correlation surface in 1846, Bravais used the same concept of two variables each composed of a large number of small elements.

12. Non-normal distributions have been classified into other types as noted in the last chapter. These may also be assumed to arise from combinations of elements, though the properties of the elements which might give rise to each type of curve have been little explored as yet. Here may lie a fruitful field of research.

13. Note that an all-or-none characteristic is the boundary between attributes and indicants. It is an attribute which is either present or absent; it is also a two-category indicant whenever the presence and absence are assigned the values of 1 and 0 ( $I^0 = 1$ ,  $I^\infty = 0$ ), respectively. When its frequency is to be counted in order to average it or treat it mathematically, it is considered as an indicant.

14. For an example of the correlation of three variables, see S. 41, Ch. VI.

15. The underlying assumption is the uniformity of nature, i.e., that conditions which in the past produced certain consequents will do so again and that variations from such a pattern are due to new factors. The whole weight of scientific experience favors this basic working hypothesis. Even, however, if it were not true as a description of apparent reality, it would still be useful as defining a frame of reference. Any deviation from past experience would still be the new factor, only in this supposition the new factors would be more capricious and the frame of reference probably less stable.

16. The coefficient of alienation  $k_{IJ}$  or its square, the coefficient of non-determination  $k_{IJ}^2 (= 1 - r_{IJ}^2)$  is a useful measure of the unknown causation of the characteristic  $I$ .  $100k^2$  gives the percent of the causes which have still to be discovered and measured before correlation (and therefore, probability and prediction) becomes perfect, evidencing complete operational understanding of the causation of the given phenomena,  $I$ . The multiple curvilinear alienation measures more generally our residual ignorance when a team of partial causes is correlated with their supposed effect. Study S. 3, 12, 16, and 22, Ch. VI in this connection.

For an example of measuring our residual ignorance see Ogburn and Talbot's study (Ref. 52) of the "factors" in the presidential elections of 1928 in the United States between Hoover and Smith. The following "factors" and their relative influence were analyzed by correlational technics:

"Factor"	Relative Influence	
Foreign born	5	The authors conclude that 59% of the "factors" of the election were measured, leaving 41% undetermined.
Rurality	32	
Democratic allegiance	663	
Catholics	109	
Wetness	314	

17. Again the definition for scientific purposes is carried only to an operational level where other scientists can check it and agreement of competent observers can be reached. Deeper metaphysical study of the essential connection, the inherent pattern, of the flux of phenomena which we observe and measure in probability and correlation coefficients and time readings, would give a fuller understanding of those phenomena without necessarily increasing our ability to predict and control them better than by operational definitions.

18. This assumes of course that the observed correlation is not a spurious or unreliable one which may arise by computational errors, such as use of ratios with common denominators, the use of inadequate samples, etc.

19. It is interesting to note that John Stuart Mills' five canons for induction of causation (the methods of agreement, of difference, of both together, of concomitant variation, and of residues) can all be quantitatively carried out by appropriate forms of correlation scattergrams involving a time differential.

20. See also S. 30, Ch. VI and S. 4, Ch. XII.

21. The advanced student will discover that these co-ordinates need not be Cartesian (i.e., orthogonal or mutually perpendicular), as oblique (i.e., non-orthogonal) co-ordinates are also possible. Orthogonal co-ordinates are simpler, however, and conform statistically to the canon of using non-overlapping categories for classifying observed data.

22. For illustration:

S. 14, Ch. III presents generalized equations re-expressing two variables,  $x$  and  $y$ , in  $n$  co-ordinates which are elemental subcomponents.

S. 23, Ch. II presents a matrix of coefficients of  $n$  indicants ("tests") analyzed into components ("factors") of three types: components common to several indicants, components specific to one indicant, and error components.

S. 34, Ch. II resolves six indicants predicting college success into two co-ordinates, namely, an indicant of that success and another indicant totally independent of it.

S. 42, Ch. VI exhibits two technics of transforming four observed characteristics into subcomponents.

23. The new journal, *Psychometrika*, is largely its organ. The sociologist interested in it but having little mathematical background cannot do better than study the introductory chapter on matrix algebra of Ref. 77.

24. A further comment or two may be of use to the advanced student in studying into the research in this field.

1. The matrix for beginning the analysis may consist of:

a. variances in the main diagonal cells and covariances in all other cells, yielding components in raw score units with a weighting dependent on the relative variances;

b. intercorrelation coefficients and, in the main diagonal cells:

i. "communalities," which yield components (in sigma units) of the common or shared parts of the indicants;

ii. reliability coefficients, which yield components of the common plus the specific part of the indicants; or,

iii. unities, which yield components of the common, specific, plus error parts of the indicants

2. A transformation of a sheaf of vectors into other vectors (or into coordinates which are but a special type of vector used as a frame of reference) may involve the following kinds of change (see S. 21, Ch. VI): translation, stretching, reflection, rotation, and distortion. Translation is moving the origin of a vector, leaving it otherwise unchanged, to the common point, 0. The amount of this translation is measured by the mean of that variable, since the origin is shifted from its observed zero to the mean as zero-point. The total amount of translation is usually the sum of the

means,  $\sum_1^n \bar{I}$  (a). Stretching is lengthening or shortening a vector,  $I$ , by

changing its scalar units, as when the variables are all re-expressed in sigma units. If the ratio of the sum of the variances of the indicants to the sum of the variances of the components equals one, no stretching has occurred, i.e., if  $\sigma_{S_1}^2/\sigma_{S_2}^2 = 1$  (b).

Reflection is reversing the sign of a variable, i.e., using its extension on the other side of the origin, 0, as graphed for  $\bar{I}_0$ , i.e.,  $-\bar{I} \rightarrow +\bar{I}$  (c). Rotation of the sheaf of vectors,  $I$ , as a whole, in each plane determined by every pair of vectors, results in reorientating the whole sheaf with respect to the co-ordinates,  $C$ , so that certain vectors may coincide with either their principal or centroidal components. The sum of the degrees of the angles of rotation divided by the number of vectors,  $\Sigma\theta/n$  (d), might serve to summarize the amount of rotation involved in a given analysis.

Finally distortion is changing the angles between the vectors, altering the angular structure within the sheaf. This distorts the facts in altering the observed correlations between the variables. A component analysis should be tested for distortion. If the determinant of the original correlations does not equal the determinant of the final components, distortion has taken place, i.e.,  $\Delta/\Delta_1 = 1$  (e) (e) is a test of distortion. (a) and (e) might be developed into percentage measures of stretching and of distortion, if needed.)

25. The following is an example of situations to which this technic for reasoning may prove valuable:

In studying competitive and co-operative behavior suppose that three variables were distinguished:

1. a competition-to-co-operation variable,  $C$
2. an economic scarcity-to-abundance variable,  $V$
3. an index of individualistic-to-collectivistic attitudes,  $I$

Suppose that conclusions were drawn out of a study which might be stated in the following propositions:

- A. People tend to compete in a situation of scarcity, and tend to co-operate in a situation of abundance.
- B. Individualists tend to compete; collectivists tend to co-operate.
- C. In situations of abundance collectivism prevails; while in scarcity, individualism tends to dominate.

To what extent are these propositions true? Are we really dealing with three

variables, or is this an illusion due to three verbal names of variables and may these be resolved into less than three variables? Is there a compulsory general component in the three named indicants?

To answer these questions some measurable indicant for each must be found and the three intercorrelation coefficients calculated. These coefficients measure the exact extent to which the "tendencies" in propositions A-C above hold good. The criterion in S. 43, Ch. VI would next show whether a general component is necessarily present. Next the optional technics would show whether two general components were sufficient to account for the observed correlations.

It is to be hoped that sociologists will increasingly reason with the aid of technics such as these and get further away from the philosophic tradition of basing propositions such as A-C on informal observation and subjective judgment in verbal terms with shifty referents.

26. The specific elements lower the correlation and produce the alienation. Their number may vary between the limits of zero and the total number of elements in the index. At the lower limit, if  $i = 0$ , index I becomes wholly a part of index J (as an addend is a part of a sum of addends) and the formula becomes

$$r_{IJ} = \sqrt{\frac{c}{c+j}} = \text{correlation of a sum (J) and a part of itself (I)} \quad (\text{Eq. 13b, Ch. VI})$$

If multiplied by 100,  $r^2$  would be the percentage that I is of J, i.e., the percentage that the number,  $c$ , of elements making up I are to the number  $(c+j)$  of elements making up J. At the upper limit,  $c$  is zero and there is no correlation,  $r = 0$ . Intermediate is the special case where  $i = j$ , i.e., where I and J have the same number of elements. Here

$$r_{IJ} = \frac{c}{c+i} = \frac{c}{c+j} = \text{correlation of two indices with the same number of elements} \quad (\text{Eq. 13c, Ch. VI})$$

Here  $100r$  is the percentage of common elements in each of the indices.

27. The more usual procedure is to assign a value of zero to its absence. If its presence and absence are equally frequent, this results in a mean of .5 and a sigma of .5. To assign  $-1$  as the value of the absent characteristic is merely shifting to deviation units, so that  $\sigma = 1$ , and shifting the mean as origin to zero.

28. The equations paralleling these verbal statements are:

$$\frac{\sum_{1}^{n/2} (+\epsilon) + \sum_{1}^{n/2} (-\epsilon)}{n} = \frac{\sum_{1}^{n} \epsilon}{n} = M_{\epsilon} = 0 \quad (\text{Eq. 14, Ch. VI})$$

$$\sigma^2 = \frac{\sum \epsilon^2}{n} = \frac{\sum (1)}{n} = 1; \quad \sqrt{\sigma_{\epsilon}^2} = \sigma_{\epsilon} = 1 \quad (\text{Eq. 15, Ch. VI})$$

$$I = \epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots \epsilon_n \quad \text{Definition of an index composed of } \epsilon \text{ elements} \quad (\text{Eq. 16, Ch. VI})$$

Since the elements are by definition uncorrelated, the variance of their sum is the sum of their variances:

$$\sigma_I^2 = \Sigma \sigma_e^2 = n \quad \text{the number of elements as the variance of an index} \quad (\text{Eq. 17, Ch. VI})$$

Next, if the two indices, I and J, are built up as the sum of the elements, the number of elements in each is given by their variances as,

$$\sigma_I^2 = c + i \quad (\text{Eq. 18a, Ch. VI})$$

$$\sigma_J^2 = c + j \quad (\text{Eq. 18b, Ch. VI})$$

Substituting these in the product moment correlation formula (Eq. 3, Ch. VI):

$$r = \frac{\Sigma IJ}{P \sigma_I \sigma_J} = \frac{\Sigma IJ}{P \sqrt{c+i} \sqrt{c+j}} \quad (\text{Eq. 19a, Ch. VI})$$

and equating this to Eq. 13, Ch. VI gives:

$$\begin{aligned} \frac{\Sigma IJ}{P} = p_{IJ} = c \quad & \text{the covariance as the number of common elements} \\ & (\text{Eq. 19b, Ch. VI}) \\ & = \sigma_I \sigma_J r_{IJ} \quad (\text{Eq. 19c, Ch. VI}) \end{aligned}$$

Equations 14 through 19 may be said to define our “ $\epsilon$ -elements hypothesis.” This is an extension to more than one variable of our “multiplex elements hypothesis,” presented in Chapter V in connection with normal frequency distributions.

29. In equations this is stated as:

$$\sigma_I^2 r_{IG}^2 + \sigma_I^2 r_{IS}^2 = \sigma_I^2 \quad (\text{Eq. 21, Ch. VI})$$

$$\text{and} \quad r_{IS}^2 = 1 - r_{IG}^2 = k_{IG}^2 \quad (\text{Eq. 22, Ch. VI})$$

30. The author is not aware of any loss of flexibility in this departure from statistical conventions, but is aware of several marked advantages—as in the use of the three descripts. Thus the symbolizing that an indicant is expressed in standard deviation units,  $\sigma I$ , needs no explanation or key to anyone knowing the standardized meanings of the class-interval descript in S-formulae. The symbols for contingency, for scatter diagrams, and for matrices of any kind are believed to be simpler and more flexible in S-theory than the conventional symbols (or absence of them as in the case of the scatter diagram).

The formulae accompanying each diagram are standard ones to be found in such a handbook as Ref. 22, except where explicit references are given to sources of less well-known formulae. A few formulae are definitional, defining the units or special conditions limiting other formulae, and their derivation from context should be obvious.

31. Note that:

$$\sigma I \equiv \sigma_I \quad (\text{Eq. 25c, Ch. VI})$$

i.e., that a distance of one sigma may be symbolized in S-notation by the index, I, with  $\sigma$  in the pointscript, or equally well by  $\sigma$  as a compound index written as a base letter, with I showing as its class script, the index for which  $\sigma$  is the standard deviation. The latter version is the conventional one in statistics.

32. For description and blueprints of half a dozen such mechanisms developed by the author to facilitate correlational research in the social sciences, see Document 1051, American Documentation Institute, care Science Service, 2101 Constitution Avenue, Washington, D.C.; for complete original manuscript remit 50 cents for microfilm form; \$3.20 for photocopies readable without optical aid.

## Chapter VII

### INTERRELATIONS, P<sup>2</sup>

#### I. THE MEANING OF INTERRELATION

##### A. Definitions

The term interrelation here denotes a relation of stimulus and response between two or more parties. It may be an active or suspended relation; it may be qualitative, quantitative, spatial, or temporal, between two persons, between two groups, or between a person and a group. A group is defined as a plurel between the members of which such interrelations exist. A group consists of interrelated persons. When the relations are dynamic ones of events, change, or actions, the term "interaction" will be used. (See Ch. X.) Interaction and static interrelations then constitute all interrelations.

Interrelations can be systematically set forth in the interrelation matrix:

$$\begin{array}{cccccc}
 \begin{array}{c} 'P \\ ''P \\ '''P \\ - \\ - \end{array} & \parallel & \begin{array}{c} 'P \\ ''(I) \\ '''(I) \\ - \\ - \end{array} & \begin{array}{c} ''P \\ '''(I) \\ ''''(I) \\ - \\ - \end{array} & \begin{array}{c} '''P \\ ''''(I) \\ '''''(I) \\ - \\ - \end{array} & \begin{array}{c} - \\ - \\ - \\ - \\ - \end{array} \\
 & & & & & \parallel & = {}^pP :: {}^P P : (I) \text{ the} \\
 & & & & & & \text{interrelation matrix} \\
 & & & & & & \text{(Eq. 1, Ch. VII)}
 \end{array}$$

Each person, represented by one row, is cross-classified against every other person represented by one column (and against himself in the main diagonal cells), and the relationship between each pair is indicated by the index in the cell. Small letter descripts may conveniently designate rows and represent the actor or expressor parties, while capital letter descripts designate columns in the full interrelation matrix and represent the same parties as recipients of transitive relations. The index may be an attribute, as in Moreno's spontaneity tests (see S. 52, Ch. X) where, when two persons face each other in an impromptu dramatization, they exhibit emotions of anger, fear, etc., in the course of the improvised dialogue. The emotion exhibited by one person, 'P, towards

another, "P, would be denoted by the attribute in their joint cell, "'I<sup>0</sup>, and the responding emotion of the latter person would be denoted by, "'I<sup>0</sup>,,. The relation between the pair may be an all-or-none (ordinal) indicant, as when each person in the group chooses or does not choose each other person as a roommate. The dyadic relation (as Von Wiese calls the relation between the two persons of a pair) may be expressed in a cardinal indicant, as in a social distance test. (See S. 13, Ch. VII.) The relation may be temporal as in the duration ( $T^{+1}$ ) of a friendship, or spatial as in an intervening geographic distance ( $L$ ).<sup>1</sup>\*

### *B. The Quantic Formula for the Interrelations, $P^2$*

The essence of complete interrelation is a cross-classification of every party with every other party. Obviously many, if not most, recorded presentations of interrelations are incomplete, in that only irregular cells or partial patterns of the whole matrix are presented. For the whole matrix for one index there are interrelations,  $P^2$  in number.  $P^2$  is the number of cells, the product of the number of rows times the number of columns.

This cross-classification is a form of multiplying every party by every other party. Since in multiplying, the exponents are added, the exponent of the population sector is 2 in the formulae for interrelations (1) through (4). The quantic digit of 2 in the quantic number  $t;i;l;2$ , then, identifies a situation composed of interrelations.  $P^2$ , denoting interrelations of every party times every other party,<sup>2</sup>† is the top stratum of blocks in S. 33, Ch. II, the three-dimensional model for visualizing the quantic classification. Within this stratum, however, the present chapter deals only with the static array,  $T^0$ , leaving interaction ( $|^{\dagger} \neq 0$ ) to Chapter X on Change.

In the preceding chapter the cross-classification of two indices in the scattergram was seen to be summarized by the correlation coefficient. The correlation coefficient was seen to be a scalar product by the rules of vectorial algebra and was, therefore, a quantic of the second degree and was so symbolized by  $I^2$ , which denotes a correlation. The cross-classification of parties does not seem to yield a scalar product. But the amount of interrelationship may be summarized in one index, such as a percentage of

\* For Eqs. 2-4, Ch. VII, see notes at end of chapter.

† For Eq. 5, Ch. VII, see notes at end of chapter.

maximal interrelation when a natural or definable maximum exists. It may prove possible to relate such a percentage to a cosine and so express the total amount of interrelationship in geometric terms as the angle between two vectors. If the vectors represent  ${}^pP_p$  and  $({}^pP_p)'$  (where priming indicates that the parties may or may not be identical with the unprimed parties), then maximal interrelation of 100% would mean that the two vectors coincided (i.e.,  $\cos \theta = 1$ ), while minimal interrelation of 0% would mean complete independence or unrelatedness.<sup>3</sup>

Another type of analysis of the interrelation matrix, which has been little explored as yet, is to analyze its variances. The variance of the whole matrix, the variance within arrays, and the variances of means of arrays may be expected to summarize much significant information. The variance is a scalar product of the indices in the cells. When these indices are populational ones, the variance involves squaring the  $P$ 's and so in this case involves  $P^2$  explicitly. When this matrix shrinks to its lower limit of one cell the variance of the  $P$  in that cell entry is simply  $P^2$  and the ordinary arithmetic square is seen to be the lower limit of the variance for a population of one plurel. In general any index raised to a power denoted by the integral exponent  $e$  is but the limiting case of the  $e$ th statistical moment in a population of one. The quantic digit then covers both statistical moments (means, variances, correlations, etc.) and ordinary powers of an index.

### C. *The Sociological Importance of Interrelations*

Interrelations of persons and groups and their dynamic subclass of interaction are becoming a major category for Sociology. Von Wiese, in his *Systematic Sociology* (Ref. 78), develops the three chief categories of Sociology as *relationship*, *plurality-patterns* (groups), and *processes*. In our quantic system the relationships are expressed in indices of relation (I). Groups are plurels having such relations between the members,  ${}^pP :: {}^pP : (I)$ , and processes are largely these with the time component included in the situation,  $T^{-1} : {}^pP :: {}^pP : (I)$ . Simmel in 1895 defined society as follows: "Society, in its broadest sense, is found wherever several individuals enter into reciprocal relations." (Ref. 63, p. 413.) Other sociologists have developed the concept. Eubank, in his *systematic Sociology* where he reviews the litera-

ture on "relationship," makes it one of seven major categories. (Ref. 25.) In our system in this volume we regard interrelations, the field of this chapter, defined by the descriptive formula Eq. 4, Ch. VII, or the quantic formula of  $\mathbb{P}^2$ , as the core of Sociology. Much of the S-theory systematizes quantitatively expressed phenomena in all the social sciences, and to that extent would be considered by many sociologists as exceeding the bounds of Sociology. But there will probably be more agreement among sociologists that the phenomena of interhuman relationships are distinctly and peculiarly the field of Sociology conceived as one of the social sciences, rather than as a synthesis of them.

These interrelations are most systematically presented in the interrelation matrix. In the first place, the matrix provides a cell for every possible interrelation which can exist. By providing columns and rows for every person and every combination of persons pertinent to any study, every possible interrelation is provided for in the appropriate cells. Too often interrelations between selected persons have been studied to the neglect of others in the situation, with no proof of the representativeness or completeness of the selection. Just as correlation, by including all cases, overcame the bias in older anecdotal technics of arguing from illustrative cases only, so the matrix, which calls attention to empty cells representing unobserved relations, is a technic tending towards completer observation of the situation and guarding against generalizing from partial data.

In the second place, the matrix concentrates attention on one characteristic at a time; it isolates for study one index interrelating all pairs of parties. Other correlated indices require further matrices—one interrelation matrix of the second degree to every index. Interrelations of persons and groups are disentangled from correlations of their characteristics. Thus, writing out the matrix is an operational technic towards achieving the scientific procedure of isolating one variable at a time, and observing its correlates, before going on to study that variable in combination with other variables.

In the third place, the matrix expresses clearly the essential property that an interrelation is not an attribute or characteristic of either party alone, but only of both jointly. It involves a pattern (Von Wiese's "plurality-pattern") of at least two persons and a relation between them. Mathematically, it is a product of two parties.

In the fourth place, the matrix is an operational definition of an interrelation. If the row and column person or group cannot be identified, the alleged interrelation is not an interrelation, but some other type of characteristic. A true interrelation must have the parties who are interrelated identified. This operational definition of constructing the matrix should reduce verbal vagueness in dealing with interrelations.

Finally, the interrelation matrix serves as a tool for further synthesizing and analyzing operations. Societal processes, as will be seen in Chapter X, can be derived as indices summarizing parts of such matrices, such as means of certain configurations of cells, sigmas of columns, correlations of rows, etc. There is the possibility, with research, of applying fruitfully, the mathematical technics of transforming matrices, depending on their rank, and analyzing out new uncorrelated components which may yield greater understanding and control of the phenomena involved.

As an example of some of these features of the interrelation matrix, Moreno's highly significant and pioneering studies may be cited. His book, *Who Shall Survive* (Ref. 48), was used by the present author, and every graph, table, or formula in it was reduced to a descriptive S-formula. (See S. 2, 3, 4, Ch. VII, for examples.) The result was, first of all, to reveal that almost every one of them had the quantic of P<sup>2</sup>. (In fact this quantic had previously been almost unused in the author's experience in analyzing sociological quantitative records.) Next it was found that Moreno's situations grouped themselves into families according to the form of the descriptive formula. From this analysis, missing members of "families" and further possible "families" suggested further data and research investigations, or experiments, which would be desirable to extend our knowledge of these phenomena. Moreno's diagramed "atoms" (see S. 24, Ch. XII) proved to be the row and column of one person in the matrix. Phenomena of leadership, popularity, etc. can be quantitatively studied by similar selection of arrays of the matrix. All the variant forms of interrelation matrices, discussed below in Section IV of this chapter, grew out of reflection upon and analyses of Moreno's data in his book and similar data in his new journal, "Sociometry," "a journal of interpersonal relations" (Ref. 65).

## II. INTERRELATIONS CLASSIFIED BY CONTENT, i.e., BY SECTORS

For a systematic treatment of interrelations, they will be classified on two bases: first, on the basis of their *content*, accordingly as they involve indicatory, population, spatial, or temporal indices; and next, on the basis of their *form*, subclassified (a) according to the chief shapes of the three-dimensional surface of the matrix, and (b) according to the form of their internal structure in combinations and permutations of their membership.

### A. *Experimental Verification of Any Classification*

Most of the classifications of interhuman relations hitherto have suffered from the defects of verbal abstraction which is not closely tied to objectively observed data. The resulting subjectivity has produced a different classification for every author.

Thus, one of the best and most thorough classifications is that of Von Wiese (Ref. 78). His organization of interhuman relations is masterly on the verbal level, but on an experimental level, leaves much to be desired. This is not a condemnation, for sociologists have not, hitherto, demanded experimental testing of the classifications of phenomena which they invent. Probably most sociologists do not conceive of how an experimental verification of a classification is possible. Yet it is a simple procedure which should become routine. In outline, it is simply for several sociologists to collect items of the kind to be classified, making either a complete collection, or one that with proper sampling controls is representative. This forces several sociologists to agree on the definition of the field to be classified and tests the workability of that definition. Next, each independently assigns each item to what he considers its proper class in the classification. This tests the inclusiveness of the classification, for, if there are items which do not fit into any one class, then the classification is proven to be incomplete. Either new classes are required to fit the facts, or else the definitions must be altered so as to include these items in some existing class, or exclude them from that classified field. Some index of agreement,<sup>4</sup> such as the percentage of discrepant assignments of items into classes, is next computed. This tests the ambiguity of the classes. The percentage of discrepancy will rise, in proportion to the overlapping of the classes. A perfect

classification is one with non-overlapping classes, and is proven by competent persons to be such by a zero percentage of discrepant allocation of items into classes.

Such experimental verification of the classification of societal phenomena as S-theory offers has been made for the quantic classification as described in Chapter II. Experimental verification of the subclassification of interrelations (the  $P^2$  class of the quantic) by content is in the first stage of gathering an adequate sample upon which to experiment. Pending the result, the classification by content will be stated here as an hypothesis which is in process of being tested.

We followed Eubank's analysis (which modifies Von Wiese's). (Ref. 25, p. 307.) He points out (giving references to Kant's classic categories) the various kinds of relationships as: <sup>5</sup>

1. spatial—our spatial index in the matrix cells,  $L^1$
2. temporal—our temporal index in the matrix cells,  $T^1$
3. similarity—our common characteristic in the matrix cells,  $I^1$
4. cause and effect—our combination of correlation,  $I^2$ , and time sequence,  $T^{-1}$ , as analyzed in the preceding chapter,  $T^{-1}I^2$
5. reciprocity (interaction)—our combination of causation and interrelations,  $T^{-1}I^2P^2$

Eubank points out that these general relationships are only societal ones, when they grow out of mental interaction. (We express this by limiting interrelations to relations of stimulus and response, which are two-way relations in the complete formula Eq. 4, Ch. VII, of our definition above.) Of these five types we shall deal first with the third kind, indicatory interrelations, and with the others in their turn, reserving those involving time indices for Chapters X and XI.

### *B. Indicatory Interrelations*

The indicatory interrelations are to a large extent synonymous with social distances—a concept which started as a technic of observing attitudes interrelating people. Eubank distinguishes three types of social distance: horizontal distance (intimacy), vertical distance (status), and lateral distance (similitude).

Horizontal social distance or intimacy is usually measured by

endorsement of one of five statements of attitude ranging from most intimate (such as: "I would marry —") to most distant (such as: "I would like to debar — from my country"). (See S. 12, Ch. II, and S. 1, 12, 13, Ch. VII.) Eubank suggests (Ref. 25, p. 326) that horizontal social distance depends upon the frequency, duration, intensity, and emotional quality of the contacts. This can be verified by correlating each of these four characteristics of the contacts with a social distance test score, and finding the multiple correlation. The multiple non-determination coefficient,  $k^2_{1,23\dots n}$ , would give the proportion of the social distance that is not determined by these four (in the specified experimental situation, of course).

Vertical social distance, or relative status, is dispersion on a valued index, as discussed in Chapter V. The relative status of persons or of plurels is measurable in sigma units above or below the mean. It is also measurable by a social distance margin.<sup>6\*</sup>

Lateral social distance is not a clear-cut concept to us. It seems partly a matter of the degree of differentiation, or absence of it, and is, therefore, covered in part by Eq. 21a, Ch. V. It is also, in part, the sympathy accompanying "similitude" and would be measurable by tests of sympathetic attitudes.

Of course, the terms "horizontal," "vertical," and "lateral," denoting three perpendicular dimensions, are figurative. They need not remain so, however, as it is only a matter of experimentally determining these three types of social distance in one population, and from their correlations determining the angles between their three (vectorial) dimensions.

Under the heading of association, Eubank summarizes the types of interaction with their resulting adjustments and status-relationships in an outline which, reproduced in part, is:

<i>Interaction Relationship</i>	<i>Adjustments</i>	<i>Status Relationship</i>
Opposition conflict competition	Elimination, $P_{,,,-0}$	Incommensurable
	Subjugation, $P_{,>,,}$	Inequality
	Compromise, $P_{,-,,}$	Near equality
Accommodation combination fusion	Alliance, $P_{,+,}$	Equality
	Integration, $P_{,,,}$	Identity

\* For Eq. 6, Ch. VII, see notes at end of chapter.

These adjustment relations between persons or between political and economic groups are approximately designated by somewhat more specific terms in personal, business, or national fields, such as:

<i>General Term</i>	<i>Between Persons</i>	<i>Between Business Firms</i>	<i>Between Nations</i>
Elimination	Murder	Bankruptcy	Belligerency
Subjugation	Slavery	Monopolistic competition	Rivalry
Compromise	Acquaintance	Competition	Independence
Alliance	Partnership	Trade association	Alliance
Integration	Intermarriage	Merger	Federation

Terms such as these are useful hypotheses, roughly indicating an interrelational continuum which has been variously described as one of approach—withdrawal, binding—loosing, association—dissociation, sympathy—antipathy. To test the hypothesis of a continuum will require devising schedule cards itemizing objectively definable attitudes, actions, and material equipment in specific fields, and refining these by appropriate statistical techniques, until scales emerge which will measure (and thereby operationally define) the interrelational continuum. Social distance tests are one form of such a measure for attitudes; but actions, events, equipment, organization of personnel, and similar phenomena will indicate interrelations even more fully. (See S. 7, 10, 11, Ch. VII; S. 50, 52, 54, 55, 56, 57, Ch. X; S. 37, Ch. XI for examples.) The measurement of interrelations of people is perhaps the major field for sociological research.

### *C. Temporal and Spatial Interrelations*

The interrelations of people may involve temporal indices in the cells of the interrelation matrix, as in situations recording the duration of friendships, ( ${}^P P :: {}^P P : T^{+1}$ ), the frequency of visits, ( ${}^P P :: {}^P P : T^0$ ), or the velocity and acceleration of any interaction, ( ${}^P P :: {}^P P : T^{-1,-2}$ ). (See S. 50, 52, 54, 55, 56, 57, Ch. X; and S. 37, Ch. XI.) The durational and dynamic interrelations are taken up in Part V, Chapters IX, X, and XI in studying time indices of societal situations, ( $T^* P^2$ ).

Spatial interrelations exist when people are separated by geographic distances and by varying degrees of accessibility, i.e., facilities of transportation and communication which reduce geographic separation. Such relations are the distances between

clienteles and their specialists in some center, such as school, church, library, playground, market, factory, or office. These distances are often better expressed in units of travel time, financial cost, frequency of transportation service, etc., than in units of physical length. (See S. 4, 7, 13, Ch. VIII; S. 84, Ch. X. The spatial relation is explicit, but the plurels interrelated are not explicit in these situations as recorded.)

Crowds are plurels with a spatial element in their definition. Fuller discussion of these phenomena will be postponed to Chapters VIII and X, when the spatial and temporal indices and their combinations in patterns of interrelated people (groups) are systematically dealt with in their turn.<sup>7</sup>

#### *D. Populational Interrelations—Mobility*

The last type of interrelations classified by the indices that occupy the cells of the interrelation matrix, is a populational one. Here a number of persons, P, is the cell entry. This usually occurs when plurels are the interrelated parties. The matrix shows the interpenetration, or overlap, of plurels. Thus, the cell entry might be the number of members common to two plurels. It might be the number of persons in one plurel which were formerly in the other, as when the foreigners in a country are cross-classified by nationality of origin, or conversely, when several nations' citizens abroad are cross-classified by country of residence.

The interdenominational exchange of church members, the interoccupational experience of employees, racial intermarriages, and all censuses of migration and mobility of population are further examples of populational interrelations.<sup>8 \*</sup>

### *III. INTERRELATIONS CLASSIFIED BY THE FORM OF THE SURFACE OF THE MATRIX*

#### *A. The Interrelation Surface*

When interrelations are classified on the basis of the surface of the interrelation matrix, operational definitions of isolation and contact, leadership and popularity, in-group and out-group, and other sociological concepts, emerge. The "interrelation" surface is the upper surface or face of the solid whose length and breadth is the number of rows and columns of the interrelation

\* For Eq. 7, Ch. VII, see notes at end of the chapter.

matrix, and whose height is a variable determined in each cell by the cell index. Thus, in the formula  ${}^{\text{P}}\text{P} :: {}^{\text{P}}\text{P} : (\text{I})$ , there will be P<sup>2</sup> values of (I). Of the shapes of this surface the chief types may be classified as "peaks," "ridges," and "plateaus" accordingly as cells, arrays of cells, or areas of adjacent cells, rise above the level of the surrounding cells. These may be referred to as "inverted" peaks, ridges, or plateaus, whenever cells, arrays, or areas, sink below the level of the surrounding matrix. These are terms borrowed from Geology, to aid in visualizing cells, arrays, or blocs of cells, which are more exactly definable as extreme deviants (i.e., beyond a specified point) from the mean.

### *B. Peaks—Points on the Interrelational Surface*

#### 1. DYADS

Whenever the interrelation of one party to another exceeds the interrelations of the other parties to each other, that one cell shows a "peak" in the interrelation surface. In a girl's boarding school, the adolescent "crush" of a younger girl on an older one is such a one-way relation of high intensity. Wherever the relation is either (a) mutual, or (b) reciprocated, there will be twin peaks in the two cells  $'::(\text{I})$  and  $''::(\text{I})$  symmetrically located with respect to the main diagonal of the matrix. This is exemplified whenever two people fall in love with each other, intensifying the index of relation in their two cells of the matrix to the exclusion of all other cells representing their other acquaintances. All mating, pairing, the "dyads," more fully discussed by Von Wiese, are instances of peaks in the interrelation surface. Degrees of difference in interrelations, critical points separating class-intervals of interrelations and defining categories of friendship, hostility, and similar relationship terms, are geometrically amounts of altitude or points on altitudes. One function of the surface is to visualize the relative interrelations; another function is to facilitate more precise observing of interrelations by expressing them as linear magnitudes.

#### 2. CLIQUES, OR BLOCS

Peaks are "pointed" when they rise in a single cell, but "blunt" peaks occur when a cluster of adjacent cells rise well above the

rest. Since the arrays are commutative, rearrangement of them can often discover or produce such blunt peaks. These are cliques sociologically, if the index is one of cohesive relations. They may vary in size from triads up, until they cease to be a "peak." Beyle has studied such cliques of interrelated legislators when the index is the number of similar votes on legislative bills, under the concept of "attribute cluster blocs." (Ref. 4. See S. 10, Ch. VII.) The matrix formulation is a valuable aid in identifying such cliques, defining major and minor cliques and "fringes" in topographic terms comparable to peaks, zones between altitude contours, etc. The matrix assists in establishing order and definiteness of classification in complex and semi-intangible patterns of sociological interrelations. The matrix assists in systematizing data; but the essential and more difficult scientific problem is still, to improve the data, to develop more adequate indices of interhuman characteristics.

### *C. Ridges—Lines on the Interrelation Surface*

#### 1. LEADERSHIP—RECTILINEAR RIDGES

A ridge in the interrelation surface is an array that is an extreme deviant from the mean of arrays (or from the mean of the matrix). A rectilinear ridge identifies the party which in that particular group of parties is outstanding. (A rectilinear ridge is that of one row or column in contrast to a diagonal array.) A single rectilinear ridge reveals a one-way relation, such as that of the "star" or *hero*, who may influence many other people without direct individualized reciprocal influence from them upon himself. A cross-rectilinear ridge, i.e., a row and the corresponding column, reveals the *leader*. The nature of the indices in the cells specifies the type of leadership, of course. But a cross-ridge indicates the person whose two-way relations to and from the other persons of the group, which is specified by any matrix under study, are highly unusual. The average altitude of the ridge measures the degree of leadership of that person, and the slope of the ridge indicates the dispersion of his leadership, i.e., the extent to which either incoming or outgoing influences vary in being stronger with some members of the group than with others. The average degree of leadership of one person can be interpreted in sigma

units from the mean of the indices of the matrix. Thus, using the terminology developed in discussing "social problems," a "maximal" would be a person the mean of whose interrelation indices exceeds some defined point, such as two sigma above the mean of all indices of the group which is operationally defined by the matrix.

$M(' :: P) > +2\sigma(I) =$  a brief formula defining leadership (of the sort indicated by the index (I)) (Eq. 8, Ch. VII)

This states that the mean,  $M|$ , of the indices of one person, 'I, cross-classified with all the persons,  $::P|$ , exceeds two sigma above the mean of all indices. If the index is normally distributed he will be one person in fifty. The outstandingness of the leadership varies with the coefficient 2. If 2 becomes 6, the leader is one person in two billion and so is unique in the living population of the world today.

At the other extreme, a "minimal" is a person characterized by an average amount of interrelation which is less than some defined point, such as minus two sigma.<sup>9\*</sup> Thus, if the relation is one of social distance, the minimal would be the hermit, the ostracized, the exiled, the isolated person; or the minimal might be the enemy, if the statements of attitude, which define any particular test of social distance, range from most friendly to most hostile.

For some examples of sociological studies of rectilinear arrays of the interrelation matrix, the "atoms" of Moreno, Lundberg, and Runner (S. 14, Ch. VII and S. 24, Ch. XII), and tests such as the ascendance-submission tests of Allport may be noted (Refs. 1 and 48). Moreno diagrams the relations of one person to all others in the group as a radial atom with nucleus and periphery (see S. 24, Ch. XII). An "atom" is the row and column of one person in the interrelation matrix. A better term is "socius," a person as a companion, i.e., interrelated to other persons. Then the interrelation matrix of P socii might be referred to as a "sociation," or more simply as a human group.

The diagram is more graphic, but the matrix form of presentation insures complete and systematic observation of interrelations with less possibility of bias through selection of unrepresentative

\* For Eqs. 9a-b, Ch. VII, see notes at end of the chapter.

cells or arrays. In terms of the interrelation surface, Moreno's "popular" vs. "friendly" persons show unequal row and column ridges. (Compare S. 12, Ch. II and S. 1, Ch. VII, for examples of unequal friendliness and popularity.) His "powerful" persons are those with only a few friends, but friends who are leaders. (See S. 3, Ch. VII.) The interrelation surface identifies these "powerful" persons by peaks in their arrays at their intersections with the ridges of those leaders. Thus, the geometric interpretation can assist in quantitatively determining these concepts of the sociological structure of a group.

One-sided relations of some people to others is still the rule in quantitative studies in the sociological literature. Very few quantitative studies of two-way relations of some people towards others, and of those others towards them, seem to have been published hitherto. Thus, the large literature of social distance tests is almost always concerned with distances of some parties towards others, and seldom ascertains the reverse distance of those other parties towards the former parties. (For two exceptions, where two-way distance was obtained under comparable conditions, see S. 12, Ch. II and S. 1, Ch. VII.)

Of course, a pair of rectilinear arrays, the "atom" of interrelationship, may be selected for study by itself, regardless of whether it is an unusual deviate, i.e., a ridge, or not. Thus, in S. 56, Ch. X, Japan's import and export relations with other nations represent, when arranged in the form of a rectangular matrix, the row and column of Japan's interrelations in foreign trade. In S. 52, Ch. X, Moreno presents the emotional attitudes of "Elsa" towards each of four other girls, and the attitudes of each of those four towards "Elsa."

## 2. SOCIETAL ISOLATION—THE MAIN DIAGONAL RIDGE

A fundamental property of the interrelation matrix for sociologists is that it provides a mathematical definition of the concepts of isolation, contact, interaction, and in-group and out-group, through the comparison of the main diagonal ridge with the rest of the matrix on either side.

Whenever the cell indices indicating the nature and amount of some positive interrelation exist only in the diagonal cells, and are zero for all other cells, the parties are isolated (with respect to the characteristics measured by those indices). Such a matrix



Ch. VII, not the social distance, but the social nearness, the complement from 100 of the percentage social distance score, should be used. The social distance scale has to be reversed. (A comparable, though not arithmetically equal, index could be used in turning Eq. 11 upside down, i.e., taking the reciprocal of  $I_1$ .)

Equations 10 and 11 can become powerful tools for more rigorous study of the sociological phenomena of isolation. They bring to this study the precision of matrix algebra. The different sociological forms of isolation, such as geographic, physiological, cultural, and attitudinal (Ref. 25, p. 318) can all be handled insofar as suitable indices can be devised to indicate such causes of isolation.

Similarly, the alleged consequences of isolation in the retardation that follows early isolation, the degeneration that follows isolation later in life, and the originality that develops from alternating isolation, can be studied more accurately, proportionately as the terms become better defined and measured. As usual, the chief methodological problem is observational, i.e., to obtain better data. The technics for handling such data have been more highly developed in advance.<sup>13</sup>

#### *D. Plateaus—Planes of the Interrelation Surface*

Logically, the study of isolation next leads to the study of contact phenomena. This also follows geometrically, since the study of isolation defined by the main diagonal ridge is completed by the study of the two triangular planes which lie on either side of the diagonal, and which define "contact."

### 1. BIPLANAR SURFACES

#### *a. Societal contact*

Contact may be defined sociologically as one-way action, in which A influences B whose responses, however, do not stimulate A in return. Movie audiences, listeners to broadcasts, and readers of printed matter are common examples. When the movie-goer, or listener, or reader, writes to the movie star, or broadcaster, or author, it becomes two-way action or interaction.

In terms of the interrelations matrix, contact may be defined as a "triangular" matrix. A triangular matrix is a particular type of matrix which is asymmetric about the main diagonal, in

that the cell entries on one side of that diagonal are zeros, and are positive on the other side.

$$\left\| \begin{array}{ccccc} - & ':(I) & ':(I) & - & - \\ 0 & - & ':(I) & - & - \\ 0 & 0 & - & - & - \\ - & - & - & - & - \\ - & - & - & - & - \\ 0 & 0 & 0 & - & - \end{array} \right\| = p \cdot (P-P') (I) = \text{the triangular matrix defining societal contact} \quad (\text{Eq. 12a, Ch. VII})^{14*}$$

The maximum number of contacts, without interaction, in a population of  $p$  parties is  $\frac{p^2 - p}{2}$ , which is the number of cells in one triangle, excluding the main diagonal cells.<sup>15</sup>

The two-way transitive relations which constitute interaction require both triangles of the matrix on both sides of the main diagonal, since, in general, the relation of A to B is not the same as the relation of B to A. Whenever the relations both ways are equal, they constitute a matrix that is called symmetric about the main diagonal. But the general case is that of asymmetric matrices of cells,  $p^2$  in number, where  $p$  is the number of persons or plurels, whichever is the unit in the population studied. Contact, then, is a special form of asymmetric matrix where one triangle has zero indices, i.e., where return relations are non-existent. Thus, living persons cannot now influence the dead, no matter how much they may be influenced by the dead.

The interrelational matrix, of course, may exhibit all degrees of asymmetry from contact as the limit of maximal asymmetry, to equality of the two triangles as the opposite limit of perfect symmetry.

The degree of asymmetry can be measured by some index derived either as a summarizing ratio, or as a difference of the interrelations in the two triangles on either side of the main diagonal. The average social distance margin is such a difference. The percentage that the mean of the indices of the cells of one triangle is to the mean of the other triangle (taking the larger average, denoted by the prime, always as the denominator) would be one measure of asymmetry.

This index becomes zero when the interrelations become exclusively contact relations, and rises to 100% when the interrelations are equal both ways, i.e., symmetric.<sup>16†</sup>

\* For Eq. 12b, Ch. VII, see notes at end of the chapter.

† For Eq. 13, Ch. VII, see notes at end of chapter.

Symmetric interrelations may be either univalent or bivalent. A univalent relation is one that needs to be measured but once, such as the miles separating two persons. The mileage has but one value, whether it is measured as distance from person A to person B, or from B to A. A bivalent relation is one that has to be measured twice, such as the social distance of person A to person B, and the other social distance of B to A. These social distances may have two values and so be bivalent. In special cases, the two social distances (or any other bivalent relations) may be equal, and so the matrix may become symmetric. As will be seen in Chapter X, interaction is usually bivalent since it involves one transitive relation going from one person to the other and the return transitive relation.

*b. One-way relations in general—the bipartite matrix*

The interrelation surface lends itself to selection of a part of itself for more intensive sociological study. Thus, the area on one side of the main diagonal may be studied alone. This area involves one-way relations (which coincide with two-way relations when the relations are univalent), but is more general than the case of contact, since the cell indices on the other side of the diagonal may not be zero. They may be unmeasured or measured and deleted in the matrix in order to concentrate attention on one set of relations. The typical pattern is a rectangular matrix selected from one side of the main diagonal of a complete interrelation matrix.

				${}^{\circ}\mathbb{P}$	${}^{\bullet}\mathbb{P}$	${}^{\vee}\mathbb{P}$	${}^{\vee\vee}\mathbb{P}$	${}^{\vee\vee\vee}\mathbb{P}$
${}^{\vee}\mathbb{P}$	X			${}^{\vee\vee} : {}^{\vee}\mathbb{P}(\mathbb{I})$	${}^{\vee\vee} : {}^{\vee}\mathbb{P}(\mathbb{I})$	${}^{\vee\vee} : {}^{\vee\vee}\mathbb{P}(\mathbb{I})$		
${}^{\vee\vee}\mathbb{P}$		X		${}^{\vee\vee\vee} : {}^{\vee}\mathbb{P}(\mathbb{I})$	${}^{\vee\vee\vee} : {}^{\vee}\mathbb{P}(\mathbb{I})$	${}^{\vee\vee\vee} : {}^{\vee\vee}\mathbb{P}(\mathbb{I})$		
${}^{\vee\vee\vee}\mathbb{P}$			X	${}^{\vee\vee\vee\vee} : {}^{\vee}\mathbb{P}(\mathbb{I})$	${}^{\vee\vee\vee\vee} : {}^{\vee}\mathbb{P}(\mathbb{I})$	${}^{\vee\vee\vee\vee} : {}^{\vee\vee}\mathbb{P}(\mathbb{I})$		
				X				
					X			
						X		
							X	

$= {}^{\vee}\mathbb{P} : {}^{\vee\vee}\mathbb{P} : (\mathbb{I})$  the bipartite matrix (Eq. 14a, Ch. VII)

This is the usual situation in social distance tests where the attitudes of a population, which responds to the test towards each of a number of plurels, are ascertained without ascertaining the reverse attitudes of those plurels to the one population. Essentially this represents relations between one set of parties and each of a second different set of parties. This matrix may, therefore, be termed the bipartite matrix. Its formula for the general case where the parties are persons, or plurels, or a mixture of both, is:

$${}^pP_p : {}^qP_q : (I) = \text{the general bipartite matrix} \\ (\text{Eq. 14b, Ch. VII})$$

where p and q denote, as usual, the number and kind of parties in the row set and the column set respectively, and where the small letter descript denotes the actor, and the capital letter descript denotes the actee, or the recipient. For illustrations of this bipartite matrix see S. 12, Ch. VII, where the social distance of one plurel of 110 Americans towards twelve other nationality plurels is presented; S. 13, Ch. VII, where characteristics are attributed to each of two students and to a plurel of 25 students by plurels of varying size of fellow students; S. 55, Ch. X, where the changing number of acquaintances of each of 16 persons is presented; and S. 54, Ch. X, presenting increases in racial friendliness of two American sex plurels towards other nationalities.

The data for the bipartite matrix are more easily secured than for the complete interrelation matrix, but the latter is sociologically far more significant just because it is complete and not a one-sided view of the facts of interrelationship. (Cf. S. 12, Ch. II and S. 1, Ch. VII.)

Still further selections of cells which are of sociological significance can be made from the bipartite matrix. Thus, the forms of the human family can be conveniently set forth in a univalent bipartite matrix with men in the rows and women in the columns. By suitable arrangement of the commutative arrays, monogamy becomes represented by the main diagonal cells. As far as marriage relations are concerned, each monogamous pair is isolated from all others. Polyandry is represented by non-zero indices in more than one cell of a row; polygyny by non-zero indices in more than one cell of a column; and group marriage by a fuller bipartite matrix. For bivalent marriage relations, the two symmetrically

located bipartite matrices from the upper right and lower left quadrants are involved. The remaining two quadrants of the full interrelation matrix of both sexes would deal with the homosexual interrelations of men with men and of women with women.

Similarly, bipartite matrices can facilitate more systematic treatment of such interrelations as those between employer and employee, buyer and seller, teacher and pupil, parent and child, specialist and client, official and citizen, doctor and patient, minister and parishioner, leader and follower, and many others.<sup>17</sup>

## 2. MONOPLANAR SURFACES—COMMUNITIES

Thus far, the discussion of areas, or planes, of the interrelation surface has dealt with a biplanar type determined by the two areas on either side of the main diagonal. The surface of an interrelation matrix, however, may give a good fit to a single plane. Usually, the extra-relations will tend to fit a plane better alone, without the intra-relations, and the tendency will be greatest when the arrays are arranged in the order of the size of the total of the indices in each array.

Such "monoplanar" surfaces, as they may be labeled, may be (a) level, (b) singly sloping, or (c) doubly sloping.<sup>18\*</sup>

A level, monoplanar surface, is one where the indices of interrelation are all equal. The most common form of this is where the indices are one kind of attribute, and are, therefore, constants of unit value. Thus, where the parties are interrelated by virtue of common membership in some group, this membership characteristic is an attribute (unless membership exists in degrees such as "applicant," "associate member," "active member," "officer," "life member," etc., in which case the attribute becomes an ordinal indicant).

The level surface means that every index tends to equal the mean index, i.e., that the standard deviation of the indices approaches zero as the limit of perfect levelness. Such level surfaces tend to occur in equalitarian fields, such as political rights in a genuine democracy, and many types of economic relations in a communistic community. In communities stratified into castes, or into any form of privileged classes, the surface of relevant indices will be either sloping or "terraced." Proportionately

\* For Eq. 15, Ch. VII, see notes at end of chapter.

as the surface approaches levelness, there is a complete "community" of interrelations (with respect to the index of that matrix).

For a singly sloping surface consider S. 12, Ch. II, the two-way social distances between religious sects, and S. 1, Ch. VII, the two-way distances between national pleurels in Syria. The dispersion of the row totals (in the last column headed "Friendliness") shows that the rows are nearly level, i.e., that the outgoing attitudes ("friendliness") are nearly equal. The dispersion of the column totals (in the bottom row headed "Popularity") shows that the columns form a slope, i.e., that the incoming attitudes ("popularity") are highly unequal. This singly sloping surface indicates a uniform trend of the social distance margins to be all positive for some groups, all near zero for other groups, and all negative for still other groups.<sup>19</sup> \*

The slope of the plane is here measured by the dispersion of the column totals. Further study of these facts may reveal that, as the surface approximates a plane, the correlation coefficients between rows,  $r_{ij}$ , approach unity (within sampling error limits). In this case, the matrix of interrelations may prove factorable, as the matrix of intercorrelations is factorable into a general interrelational component and specific components for each party. The interpretation of such a condition, whenever its existence may be demonstrated, would be that a common interrelational set of attitudes is diffused through all the parties in the population studied. A common set of attitudes means a general attitude held by all parties towards each of the parties in turn, i.e., all parties here dislike Jews (Column 4, S. 12, Ch. II), all like Protestants (Column 5, S. 12, Ch. II), etc.

The rich development of analysis into components, that is going on from intercorrelational matrices, as sketched in Chapter VI, may become duplicated for interrelational matrices. It is conceivable that interrelational epsilon elements, as described in the previous chapter, may become determinable.

With the progress of research, such epsilon elements may prove to be as valuable units for Sociology, as the analysis of matter into atomic units has proved for Physics and Chemistry. At present this is an untried hypothesis, but it is an hypothesis which

\* For Eqs. 16, 17a-b, Ch. VII, see notes at end of the chapter.

can be definitely tested by first working out the theoretic statistical points involved, and then seeing the extent to which interrelational matrices, composed of indices of attitudes, may fit that body of statistical theory.

Monoplanar interrelation surfaces in general may be used to define a "community." A level surface is a perfect community, since all interrelations are in common and exist to the same degree for all. A singly sloping surface means a community of one-way relations with a dispersion of the reverse relations. The degree of community is defined as the percentage of interrelational value characteristics common to a group. Suppose a catalogue were made of all the kinds of interrelating indices in a group, and that these indices were weighted in proportion to their value to the group studied, i.e., in proportion to the average relative intensity of desire for each characteristic as represented by its index; suppose that the indices were determinable for every pair of members of the group and could be expressed as a percentage of their maxima, then a summarizing percent of the observed amount of interrelations to the maximum amount representing perfect community would be an index of the degree of community. "A community" would thus be operationally defined in terms of:

- a. a specified list of interrelating characteristics and indices indicating them,
- b. a specified weighting for combining these indices; a weighting operationally determined by some technic of a vote, questionnaire, time budget, or other device for objectively observing people's desires (evaluations),
- c. observed square matrices of such indices—a second-degree matrix for every index, a row and column in each matrix for every member of the group,
- d. a specified maximum for each index, either a natural, or some conventional, limit,
- e. a formula defining the calculation of a summarizing index of community from the data of a. to d.

Consider a very much simplified illustration of this quantitative definition of a community, since data for a complete illustration have not yet been collected by sociologists. In S. 12, Ch. II, we have a complete observed interrelational square matrix for the

characteristic social distance, in a population composed of six sectarian groups. Specifications a. and b. above are simplified to a list of only one index, so that no relative evaluations for weighting indices are needed. Specification c. is S. 12, Ch. II, as recorded. Specification d. is included in the index of social distance, as the index is a percentage of the maximal distance. For specification e., since the indices are percents of their maxima, a simple mean summarizes them. The mean social distance is 41%.

Reversing the scale, the mean social "nearness," or friendliness, is 59% of the maximum friendliness. This population then shows 59% of community of intersectarian attitudes as defined by this Beirut test. It may be spoken of as an intersectarian community which, *in this one respect as measured*, is 59% of a perfect intersectarian community. ("Perfect" connotes no value judgment here but only 100% on the defined scale.)

Although quantitative data are, as yet, inadequate for precise determination, this definition of a community may be conceptually applied to the list of communities in Chapter III, and Chapter IV.<sup>20</sup>

### 3. POLYPLANAR SURFACES

In addition to monoplanar interrelation surfaces useful in defining communities, and biplanar surfaces useful in defining isolation and contact, surfaces analyzable into more than two planes may be distinguishable as more interrelational matrices become reported in the sociological literature. Data which a priori seem likely to yield such surfaces, perhaps in the form of zones suggested by S. 10, Ch. VII, are such interrelations as:

- a. kinship groups—married couples, siblings, cousins, which may show decreasing indices of certain types (see S. 11, Ch. IX),
- b. primary vs. secondary groupings of one population,
- c. person-group matrices where "terraces" of significant increases in the indices may be expected in passing from the person-person arrays to the person-group arrays, and again in passing on to the group-group arrays,
- d. any inter-party relations where the parties are classified into plurels with differential barriers between plurels. Thus,

business firms in trade relations with other firms in their own country and in other countries with varying tariff walls, may be expected to show "terraced" interrelation surfaces where the height of the terraces varies with the height of the tariff wall.

Finally it should be noted that a suitable form of a goodness-of-fit index can be derived to measure the degree to which (or the probability with which) an observed surface approaches any defined type of surface. The computation of such goodness-of-fit may be carried out in order to classify observed matrices into standard types; or in order to test an hypothesis that interrelations under certain specified conditions will be of a certain type of surface; or in order to modify the indices until they conform to some experimental desideratum, such as a situation of pure contact without interaction, or a level surface reflecting an equalitarian group.

#### IV. INTERRELATIONS CLASSIFIED BY THE FORM OF INTERNAL STRUCTURE—THE COMBINATIONS AND PERMUTATIONS OF THE MEMBERS OF THE GROUP

A promising method, little developed as yet, for systematically exploring the possible structures of human groups, is with the aid of the theorems of combinations and permutations. Two cases of this method will be sketched—a case of combinations from contacts and a case of combinations from interrelation.

##### A. Groups Analyzed by Contact Combinations

Karl Menger in an article entitled "An Exact Theory of Social Groups and Relations" (see S. 8, Ch. VII) analyzes two plurels,  $P_1$  and  $P_2$  ( $G_1$  and  $G_2$  in his notation) each of which subdivides into subplurels according to their willingness to associate with each other thus:

$P_{1:12}$  = a subplurel of  $P_1$  willing to associate with both  $P_1$  and  $P_2$

$P_{1:1}$  = a subplurel of  $P_1$  willing to associate with  $P_1$  only (in-group association)

$P_{1:2}$  = a subplurel of  $P_1$  willing to associate with  $P_2$  only (out-group association)

$P_{1:0}$  = a subplurel of  $P_1$  willing to associate with neither plurel  
(the dissociative case of hermits, not diagramed by Menger  
in S. 8, Ch. VII)

$P_{2:12}$  = a subplurel of  $P_2$  willing to associate with either  $P_1$  or  $P_2$

$P_{2:1}$  = a subplurel of  $P_2$  willing to associate with  $P_1$  only

$P_{2:2}$  = a subplurel of  $P_2$  willing to associate with  $P_2$  only

$P_{2:0}$  = a subplurel of  $P_2$  willing to associate with neither plurel

For practical predictive purposes examples may be studied of smokers and non-smokers in a train compartment, couple dancing with partners of opposite sexes only, and bilingual groups knowing or not knowing the other's language. The scheme above illustrates all the possible combinations when the number of plurels,  $|_p|$ , is 2, and when their interrelating attitudes are all-or-none ones, so that the number of positive degrees,  $|_i|$ , of the interrelation is 1. The combinations are given by the binomial expansion  $(a + b)^2$ .

This case can be generalized to any number,  $p$ , of plurels and to any number,  $i$ , of either degrees, or types, of interrelation (i.e.,  $i = |_i|$  or  $|_i|$  or  $|_i|$ ) by means of the expansion of the binomial. The total number of combinations of one plurel with all the plurels in the situation is given by the sum of the coefficients of the expanded binomial. This sum is:

$$2^p = \text{total number of subplurels of one plurel, } P_{p': 2q} \\ (\text{Eq. 18a, Ch. VII})$$

As each plurel in the situation subdivides into subplurels the total number of subplurels is multiplied by  $p$ . This number,  $p2^p$ , assumes that the number of positive interrelations is one ( $i = 1$ ). But if  $i > 1$  this multiplies the number of subplurels by  $i$ . Thus if people "will associate" or "will associate on condition that . . ." or "will not associate" the number of positive interrelations is two ( $i = 2$ ). The formula then becomes

$$pi2^p = q_1 = \text{total number of subplurels of all plurels, } P_{p: 2q} \\ \text{for } i \text{ positive contact relations} \\ (\text{Eq. 18b, Ch. VII})$$

For Menger's case above, the binomial terms give  $p = 2$  and  $i = 1$  and  $q_1 = 8$  as listed above. For a more complicated instance, consider a community with four language-speaking plurels whose subplurels are determined by their ability to communicate

with each other. Let language ability be observed in three degrees ( $i = 2$ )—ignorance, reading knowledge, reading and speaking knowledge. Substituting  $p = 4$  and  $i = 2$  of this situation into Eq. 18b above gives  $2 \cdot 2^4 = 128$  possible subplurels. Some other values of  $q$  are:

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
$p = 1$	$q_1 = 2$	$q_1 = 4$	$q_1 = 6$	$q_1 = 8$	$q_1 = 10$
$p = 2$	$= 8$	$= 16$	$= 24$	$= 32$	$= 40$
$p = 3$	$= 24$	$= 48$	$= 72$	$= 96$	$= 120$
$p = 4$	$= 64$	$= 128$	$= 192$	$= 254$	$= 320$

In case hermits are precluded in the situation as specified, then 1 must be subtracted from Eq. 18a before multiplying it by  $p_i$  to get Eq. 18b. Thus, omitting hermits, Menger's case is  $2 \cdot 1(2^2 - 1) = 6$ .

The first row for  $p = 1$  expresses the self-attitudes, i.e., the number of subplurels of one plurel determined by in-group attitudes and the same  $i$  attitudes to hermits.

As the number of subplurels grows larger, they may be summarized by coarser classifying, or by appropriate summarizing indices. Thus, the community above may have its structure of 128 possible subplurels simplified by recording the four overlapping subplurels which understand each of the four languages.

The detailed structure of these groups is expressible in an interrelation matrix of order  $p \times 2^p$  by the simple arrangement of using a row for every plurel and a column for every combination of plurels from 0, 1, 2, 3—up to  $p$  plurels at a time. This means  $p$  rows and  $2^p$  columns. The cells provide an entry for one type of interrelation between every person and every person and group of persons including in-relations. The matrix presentation has the advantage over Menger's diagram above in that it can portray without confusion: (a) every interrelation in quantitative degrees (not merely in all-or-none degrees as in his diagram); and (b) every interrelation of *any* number of plurels.

### *B. Groups Analyzed by Interrelational Combinations*

The case just discussed dealt with plurels subdividing on the basis of contacts with other plurels. These were one-way rela-

tions of what the subdividing plurel wanted with no regard to any reciprocating attitudes of the other plurels. The case of two-way relations or interrelations may now be sketched.

Suppose two persons express their attitudes each to the other in the three degrees of "I am attracted, indifferent, repelled." Representing these degrees by +1, 0, and -1, respectively, the possible structure of interrelations in this pair is shown in the following scheme:

P	+	-	0	+	-	+	0	-	0	where ++ means a
P	+	-	0	-	+	0	+	0	-	pair showing mutual
										attraction, -- means
										mutual repulsion, etc.

Here there are six types of structures or combinations of +, 0, and - and nine permutations, or arrangements of the persons. Let this be generalized to any number, p, of persons or plurels, and to interrelations in any number, i, either of degrees or of qualitative types. How many different group structures, i.e., how many combinations, are possible? The answer is readily found with the help of the standard theorems of combinations and permutations.<sup>21</sup>\*

Another question is as to the number of permutations of persons in these group structures, i.e., how many patterns are there like the nine possible pairs of +, 0, -, tabled above? This number is given by the sum of the coefficients of the expanded multinomial of i terms when raised to the p<sup>2</sup> power.<sup>22</sup>†

Note that these formulae state the total possible number of subplurels. Actually in any given situation there will be only one observed degree of interrelation in every pair of parties (a party = a person, or a plurel) so that the matrix (Eq. 21, Ch. VII) can record *all actual* interrelations very simply and completely with no confusion, no matter how large p and i may be. This matrix can be summarized in a single mean, or in some other desired index, either as a whole or subdivided into sections for some special analyses.

When the interrelations are cohesive ones such as inter-person

\* For Eq. 19, Ch. VII, see notes at end of chapter.

† For Eqs. 20-21, Ch. VII, see notes at end of chapter.

attitudes of attraction-repulsion, a summarizing index such as a mean may measure the total *cohesion* of the group. Moreno's study uses such indices. Such cohesion may be expected to correlate with the social longevity of the group and to correlate negatively with turnover of membership. The higher the correlation the more exclusively are such attitudes the causes of group longevity or of minimal turnover. Here is an important sociological field of hypotheses awaiting exploration.

These brief explorations suggest a rich field for further research in systematically mapping out the possible structures of groups under varying but specified conditions, and simplifying and summarizing these for sociological applications. These explorations also suggest the basic importance of the interrelation matrix, and back of it the S-theory formula, in that it proves so flexibly able to subsume as special cases quantitatively developed formulae or theories in Sociology such as Menger's theory, and then extend them with greater generality through the standardized symbols of S-theory.

#### V. S-SITUATIONS

Although interrelations would seem to be the essence of the field of sociology yet explicitly quantified interrelational data are very scarce and patchy. Most data are mere diagrams or one-way social distance or other attitude tests in all-or-none or ordinal units. The samples of interrelational data appended here are a large proportion of all that our canvass of the sociological literature revealed—in contrast to the samples at the end of the other chapters which represent minute proportions of the published data. Only two tables were found presenting two-way interplurel relations in cardinal units with ascertained reliability. Even these two S-situations from the Beirut test (S. 12, Ch. II and S. 1, Ch. VII) fall far short of the best experimental standards that are known to date.

The fourteen S-situations that follow will, however, illustrate interrelational data, defined by the quantic  $P^2$ , as reported in the current literature. As usual, for reading and interpreting the descriptive formula and its legend and the quantic number, the explanation preceding the S-situations at the end of Chapter II may be read again.

## S. 1

SOCIAL DISTANCES BETWEEN NATIONAL AND RACIAL GROUPS.  
STUDENTS OF THE AMERICAN UNIVERSITY OF BEIRUT, SYRIA

Responders		Respondees															
	N*	Armenian	Egyptian	Iraki	Jew	Palestinian	Syrian	American	British	Chinese	French	Greek	Italian	Kurd	Negro	Turk	Weighted Average "Friendliness"
Armenian	20	(.31)†	1.6	1.8	1.1	1.4	1.5	.8	1	1.8	1.6	1.4	1.4	2.2	1.9	3.3	1.6
Egyptian	20	.8	(.3)	2	2.8	1.4	1.3	1	2.3	1	1.2	2	1.7	2.8	1.7	1.5	1.6
Iraki	43	2.4	1.3	(.16)	2.8	1.6	1.3	.8	.5	2.1	1.7	1.5	1.8	2.2	1.3	1	1.6
Jew	20	1.7	2	2	(.22)	1	2	.78	1.5	2	2.1	2	1.2	1.7	1.5	2	1.6
Palestinian	53	2.6	1.2	1.8	3	(.43)	1.2	.98	1.8	1.9	2.1	2	1.2	1.7	1.5	2	1.7
Syrian	85	2.5	1.2	1.7	2.3	1.1	(.3)	.88	1.4	2	1.6	1.5	2.1	2.5	2.1	1.4	1.7
Weighted Average "Popularity"	241	2.2	1.4	1.8	2.4	1.3	1.4	.86	1.5	1.9	1.6	1.6	1.7	2.3	1.7	1.6	1.66

\* N is the number of responses.

† Distances in parentheses are of in-groups. Their weighted average is .3. All groups of responders which were smaller than 10 were omitted in this table. The scale runs from 0 to 4.

Ref.: Dodd, Stuart Carter, "A Social Distance Test in the Near East," *Department of Sociology Year Book* (Typescript), American Univ. of Beirut, Vol. VI, Table 3, 1933. Summarized in "A Social Distance Test for the Near East," *American Journal Sociology*, Vol. XLI, No. 2, Sept., 1935.Descriptive formula:  $S_1 = P_p :: \underline{P}_{p+Q} : {}^m I$ 

Quantic number = 0;1;0;2

Legend:

 $S_1$  = The situation  
records $P_p$  = 6 National plurels $::$  = cross-classified  
with $\underline{P}_{p+Q}$  = themselves and 9 other  
plurels

: = with their corresponding

 $I$  = social distances ${}^m |$  = and mean attitudes are also  
stated

Comment:

For the first 6 plurels,  $|_p$ , the matrix is a full two-way one; but for the 9 other plurels,  $|_q$ , the matrix is a one-way or bipartite matrix. This is believed to be the first record in the sociological literature of two-way social distances between all of a set of plurels (6) measured by an attitude test in calibrated cardinal units.

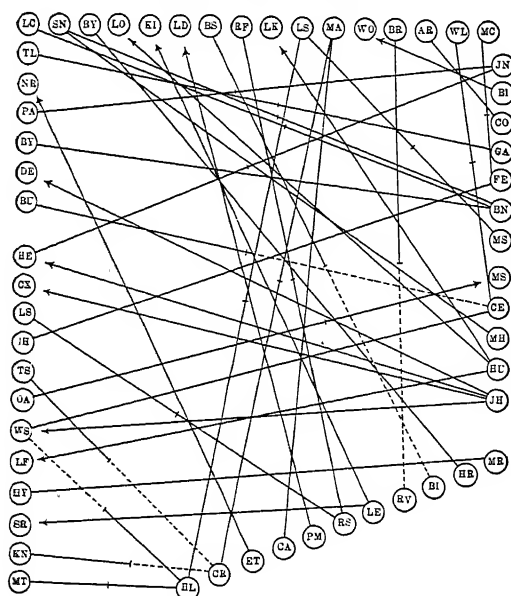
It is a singly sloping monoplanar surface, as evidenced by the constancy of the six averages in the last column (showing the rows to be of the same level) and the variability of the fifteen averages in the bottom row (showing the columns to vary in level). On rearranging the order of the columns (since these are com-

mutative), the trend for the cell entries to decrease from left to right in a plane sloping downwards to the right-hand end of the matrix would be more evident.

This singly sloping plane is the geometric portrayal of the striking sociological finding of constant friendliness (outgoing attitudes) among these plurels, in spite of highly variable popularity (incoming attitudes). The intensely disliked Jews did not reciprocate hatred on the average, any more than the more popular Palestinian Arabs did. (These findings were corroborated on two other samples of 200 upper classmen and 200 business people of Beirut. The tests were calibrated by Thurstone's technic to convert the usual ordinal units of social distance tests into cardinal units. The repetition reliability after a month's interval was  $r = .91$ . Cf. S. 12, Ch. II.)

## S. 2

### CLEAVAGES IN GROUPS



STRUCTURE OF INTERRELATIONS BETWEEN TWO  
RACIAL GROUPS

Criterion: Living in Proximity

The chart indicates the interchanged attitudes in respect to living together in the same cottage of 23 colored girls and 34 white girls. Six white and six colored girls form mutual pairs, wanting to be in the same cottage together. Fourteen incompatible pairs are formed, one party wanting to live in the same house, the other not. Seven colored girls are attracted to white girls who do not respond. The population in which the sociometric test revealed these interrelations to exist consisted of 377 white and 58 colored girls.

Ref.: Moreno, J. L., *Who Shall Survive? Nervous and Mental Disease Publishing Company, Washington, D. C., 1934, p. 214.*

Descriptive formula:  $S_2 = {}^pP :: {}^aP : {}^iI$

Quantic number = 0;1;0;2

Legend:

$S_2$  = The situation  
records

${}^aP$  = 23 colored girls  
by

${}^pP$  = 34 white girls

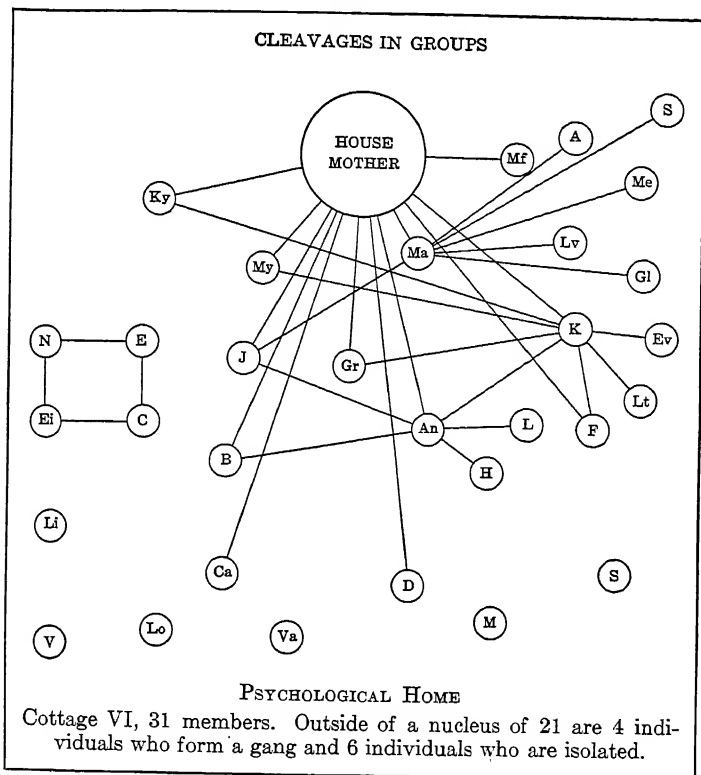
$::$  = interrelated  
with

${}^iI$  = attitudes of  
3 degrees { attraction  
indifference  
repulsion

Comment:

The original lines are in red and black and differentiate the attitudes better than their reproduction in black does here.

### S. 3



*Descriptive formula:*  $S_3 = {}^1P_3 :: {}^2P : {}^1I$

*Quantic number* = 0;1;0;2

*Legend:*

$S_3$  = The situation

${}^2P$  = each other

records

by

${}^1P$  = 31 girls in Cottage VI

${}^1I$  = attitudes of attraction or neutrality

'.' = and their housemother

:: = interrelated with

*Comment:*

Note the "ridges" of leadership. In the interrelation surface, out of  $32 \times 32 = 1024$  cells, only 64 show positive attitudes and the rest are empty (i.e., relations of indifference). This is an average of two cells per array, or two friends for each person. Yet the housemother has 12 friends "K" has 8, "Ma" has 7, and "Ak" has 6.

*Comment on notation:*

As usual the descriptive formula may be written, if desired, with numbers partially replacing letters, thus:

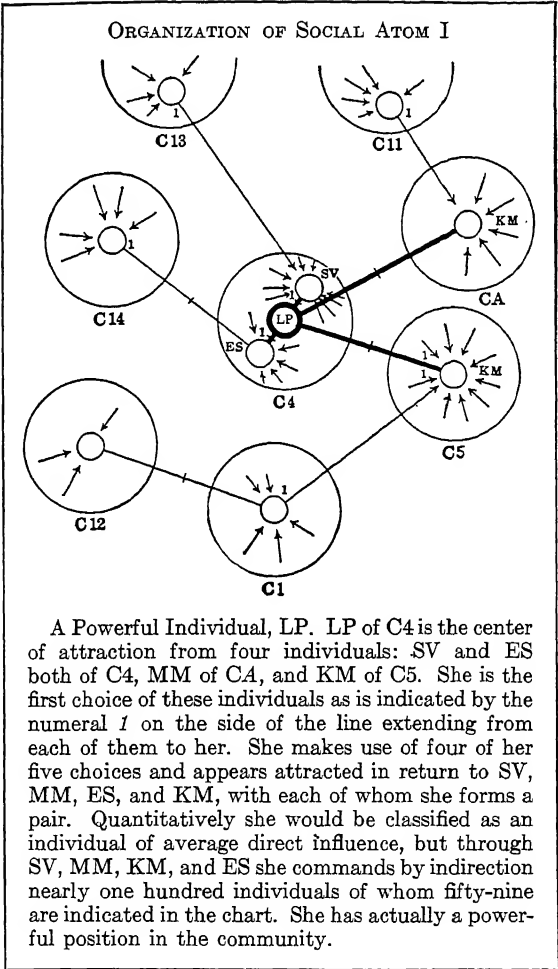
$$S_3 = {}^{31}P :: {}^{31}P : {}^1I$$

The legend can then be simplified to:

P = girls in Cottage VI

I = attraction attitudes.

S. 4



Ref.: Moreno, J. L., *Who Shall Survive?* Nervous and Mental Disease Publishing Company, Washington, D. C., p. 90.

Descriptive formula:  $S_4 = {}^{\circ}P_{,p} :: P+{}^qP : 1,0I$       Quantic number = 0;1;0;2

Legend:

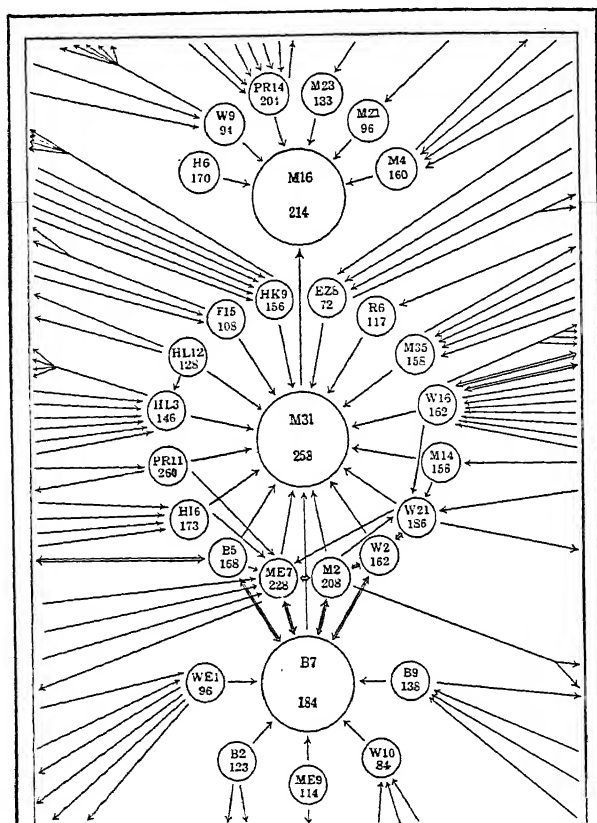
- $S_4$  = The situation  
records  
 ${}^{\circ}P$  = 9 persons  
and  
 ${}^{\circ}I$  = a "leader"  
grouped into  
 $|_{,p}$  = 8 cottages

interrelated with  
 $P|$  = each other  
 $q|$  = and 50 other persons  
in  
 $1,0I$  = attitudes of attraction or neutrality

*Comment:*

This is the social atom of a person ("LP") who is powerful through friendship with four leaders (H.M., K.M., S.V., and E.S.). In the interrelation surface the two arrays for "LP" show peaks where they cross the eight ridges of the four leaders.

## S. 5



"Friendship" Constellation in a Village. Each person is represented by a circle. The letter and the first number in the circle is the code symbol of the person. The second number is that person's score of socio-economic status on the Chapin scale. Each arrow represents a choice made or received according to the direction of the arrow. Mutual choices are represented by double-headed arrows.

*Descriptive formula:*  $S_5 = {}^pP :: {}^pP : I_{,,,}$

*Quantic number* = 0 ; 1 ; 0 ; 2

*Legend:*

$S_5$  = The situation  
records

${}^pP$  = for each of 30 persons

$::$  = cross-classified  
with

${}^pP$  = each other

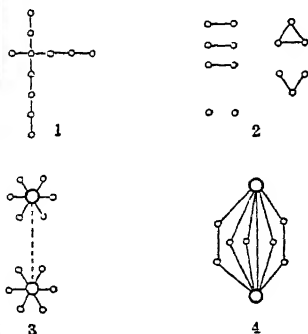
$I_{,,}$  = an indicator of choice (choosing or not-choosing, <sup>1,0</sup> $I$ )

and

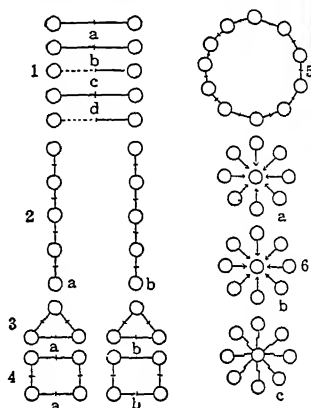
$I_{,,,}$  = a socio-economic status score  
is also stated.

## S. 6

## ORGANIZATION OF GROUPS-HOMES



## ORGANIZATION OF GROUPS-HOMES



## TYPICAL STRUCTURES WITHIN GROUPS

1. Attractions between individuals take the form of a chain.
2. Attractions take the form of isolated units, pairs and groups of three.
3. Two sub-groups are centralized each about two dominating individuals who have no attractive forces uniting them.
4. A group in which two dominating individuals are strongly united both directly and indirectly through other individuals.

## TYPICAL STRUCTURES WITHIN GROUPS

1. Attractions and repulsions take the form of a *pair*: in a mutual attraction (red pair); b, mutual rejection (black pair); c, mutual indifference; d, attraction vs. rejection; e, attraction vs. indifference.
2. Mutual attractions and mutual repulsions take the form of a *chain*: a, chain of mutual attractions; b, chain of mutual repulsions.
3. Mutual attractions and repulsions take the form of a *triangle*: a, triangle formed by attractions; b, triangle formed by repulsions.
4. Mutual attractions and repulsions take the form of a *square*: a, square formed by attractions; b, square formed by repulsions.
5. Mutual attractions take the form of a *circle*.
6. Mutual attractions and repulsions take the form of a *center (star)*: a, center of attractions; b, center of repulsions; c, center of incompatible repulsions vs. attractions.

Descriptive formula:  $S_6 = ({}^pP :: {}^pP : {}^1I)_p$

Quantic number = 0;1;0;2

Legend:

$S_6$  = The situation

presents

${}^pP$  = persons

:: = interrelated with

${}^pP$  = each other

by

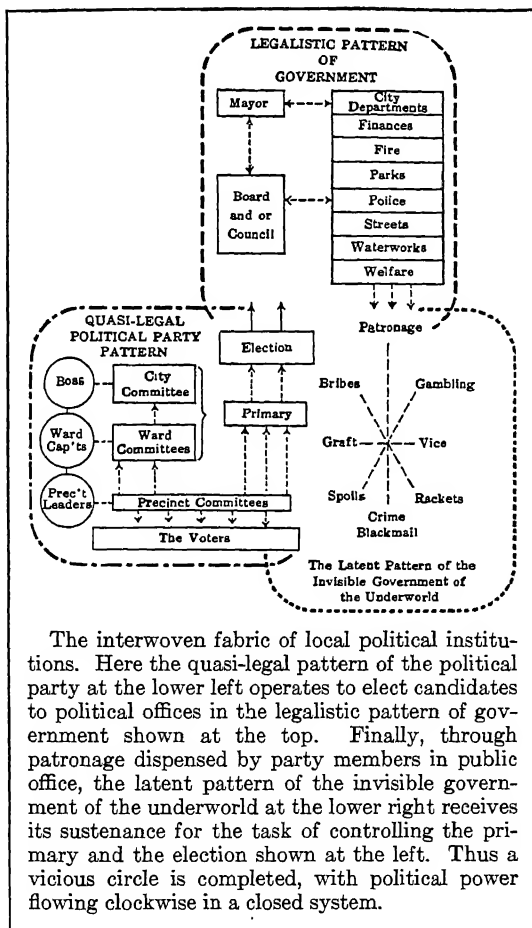
${}^1I$  = attitudes in 3 degrees

{ attraction  
indifference  
repulsion

in

$|_p$  = 22 groups or types of structure

## S. 7



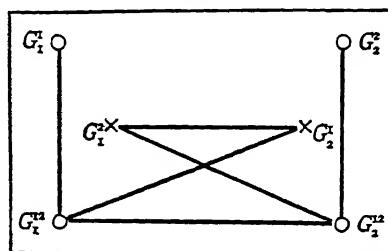
*Descriptive formula:*  $S_7 = \langle \underline{P}_{p:q} : (\cdot \underline{P})_{r:s:t:u} : I_1^0 : j \rangle$  *Quantic number* = 0;0;0;2  
*Legend:*

$S_7$ = The situation	$ _s$ = the ward committee
records	and
$\cdot \underline{P}$ = the Mayor	$ _t$ = the precinct committee
and	and
$ _p$ = Government Council	$ _u$ = the voters
and	each with
$ _q$ = 7 City Departments	$\cdot  $ = their leader
related to	in relationships of
$ _r$ = the city committee	$I_1^0$ = election and patronage
and	$ _j$ = of 7 subtypes

*Comment:*

The situation, as Chapin points out, is analyzed to show a complexity of interrelations where the analysis is not clear-cut and quantified. Hence, to symbolize it in an S-formula by the aid of matrix algebra is difficult. The formula involves more interpretation and chance for the subjective judgment of the analyst to result in disagreement than most situations. It is included as a border line case of a quantifiable recorded situation, which is the datum of S-theory.

S. 8



*Ref.:* Menger, Karl, "Our Exact Theory of Social Groups and Relations," American Journal of Sociology, Vol. XLIII, No. 5, March, 1938, p. 793.

*Descriptive formula:*  $S_8 = \underline{P}_p :: \underline{P}_q : \underline{I}^0$  *Quantic number* = 0;0;0;2  
*Legend:*

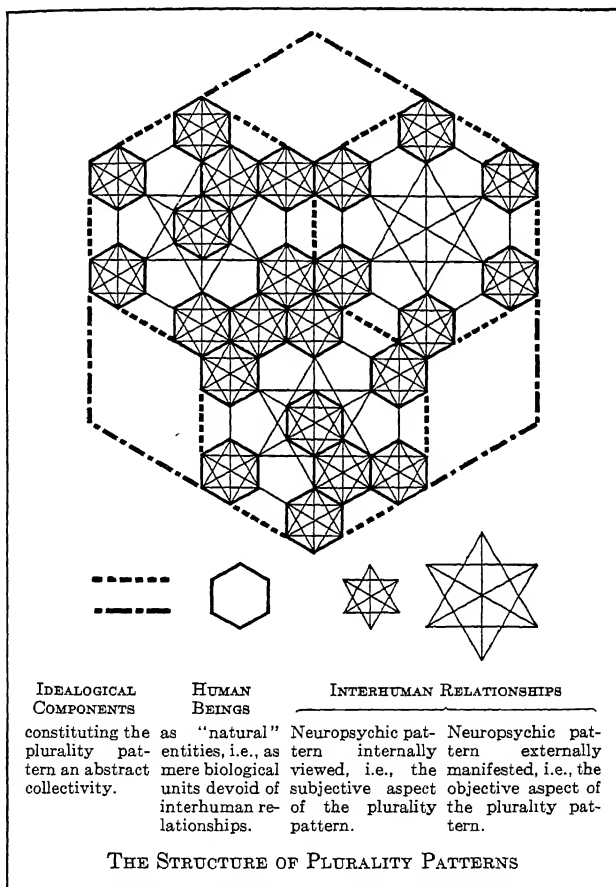
$S_8$ = The situation	$\underline{P}_q$ = 3 other plurels ( $G_1^1 G_1^{12} G_2^2$ )
records	by
$\underline{P}_p$ = 3 plurels ( $G_1^1 G_1^{12} G_1^2$ )	$\underline{I}^0$ = an unspecified relationship
:: = interrelated with	(symbolized by a connecting line)

*Comment:*

The unlabeled diagram represents the subdividing of each of two plurels ( $G_1$  and  $G_2$ ) according to their willingness to associate with their own plurel,

with the other plurel, or with both plurels. Thus, if  $G_1 = \text{men}$ , and  $G_2 = \text{wome}$  and the relationship is couple dancing, most of the population would be four in  $G_1^2$  and  $G_2^1$ , since people prefer to dance with a partner of the opposite sex (The equivalent S-notation is  $P, + P,,, \text{ and } P, :,, P,,, :,,$ ) This diagram explores the total possible combinations of associating between two plurels as a special case of a more general mathematical theory of groups which is developed at the end of Chapter VII.

## S. 9



*Descriptive formula:*  $S_9 = {}^{\text{P}}\text{P}_{\text{, : : } q} :: ({}^{\text{P}}\text{P} : \underline{\text{I}}_1^0) : \underline{\text{I}}_2^0$       *Quantic number* = 0;0;0;2

*Legend:*

$S_9$  = The situation

diagrams

${}^{\text{P}}\underline{\text{P}}$  = 23 unidentified persons

[who are stated to be in

$|_{\text{,}}$  = one plurel (i.e., the "abstract collectivity," marked "-.-.-")

subclassified into

$|_q$  = 3 subplurels ("—")]

$::$  = interrelated with

${}^{\text{P}}\text{P}$  = the same persons

(each having

$\underline{\text{I}}_1^0$  = "internal neuropsychic-patterns")

by means of

$\underline{\text{I}}_2^0$  = "external neuropsychic patterns"

*Comment:*

Two of Von Wiese's three categories of plurality patterns and their equivalents in S-theory are shown here. His "abstract collectivity" of human beings, Eubank's "aggregation," is a plurality of persons with some characteristic in common. This is denoted by our class script, which is the plurel script on the population index,  $|_{\text{p,q,etc.}}$ . Von Wiese's and Eubank's and the author's concept of "a group," which is a plurel of interacting members, is represented in summary by the quantic of  $\text{P}^2$  and in detail by the qualitative interrelations,  $\text{I}^0, j$  in number, between the persons specifying the structure of the group. The attributes  $\text{I}_1^0$  and  $\text{I}_2^0$  may be expected to be of low reliability as they are specified by verbal names ("neuropsychic patterns") and not by operational definitions of any schedule card or scale.

## S. 10

TABULAR WORK SHEET NO. 8a. TABULAR TREATMENT OF FIRST BLOC SYSTEM										
Principal bloc										
Nucleus				Inner fringe of adherents			Outer Fringe of adherents			
	6	1	9	2	3	4	5	7	8	
6		33.3	33.3	33.3	16.6	33.3	16.6	16.6	—	
1	33.3		33.3	16.6	16.6	16.6	—	—	—	
9	33.3	33.3		16.6	16.6	—	—	—	—	
2	33.3	16.6	16.6		33.3	16.6	33.3	33.3	33.3	
3	16.6	16.6	16.6	33.3		16.6	16.6	16.6	16.6	
4	33.3	16.6	—	16.6	16.6		16.6	16.6	16.6	
5	16.6	—	—	33.3	16.6	16.6		33.3	16.6	
7	16.6	—	—	33.3	16.6	16.6	33.3		16.6	
8	—	—	—	33.3	16.6	16.6	16.6	16.6		

Items are indexes of significance of cohesion of pairs on high scores.

Indications of existence of other bloc systems

TABULAR WORK SHEET NO. 8b. TABULAR TREATMENT OF SECOND BLOC SYSTEM										
Principal bloc										
Nucleus				Fringe of adherents						
	2	5	7	3	6	4	8	1	9	
2		33.3	33.3	33.3	33.3	16.6	33.3	16.6	16.6	
5	33.3		33.3	16.6	16.6	16.6	16.6	—	—	
7	33.3	33.3		16.6	16.6	16.6	16.6	—	—	
3	33.3	16.6	16.6		16.6	16.6	16.6	16.6	16.6	
6	33.3	16.6	16.6	16.6		33.3	—	33.3	33.3	
4	16.6	16.6	16.6	16.6	33.3		16.6	16.6	—	
8	33.3	16.6	16.6	16.6	—	16.6		—	—	
1	16.6	—	—	16.6	33.3	16.6	—		33.3	
9	16.6	—	—	16.6	33.3	—	—	33.3		

Indication of first bloc system

	5	7	8
5		33.3	16.6
7	33.3		16.6
8	16.6	16.6	

On treatment of data from high end of range of scores, bloc appears as a lesser system. Same bloc appears as only important system when data from low end of range of scores are treated.

Ref.: Beyle, Herman C., *Identification and Analysis of Attribute-Cluster-Blocs*, The University of Chicago Press, 1931, p. 59.

Descriptive formula:  $S_{10} = ({}^P P_{.a} :: {}^P P : I)_P$

Legend:

$S_{10}$  = The situation

records

${}^P P$  = 9 persons

$_{.a}$  = grouped in 4 plurels

{ nucleus  
prin. bloc  
inner fringe  
outer fringe

$::$  = cross-classified with

Quantic number = 0;1;0;2

${}^P P$  = the same 9 persons

$:$  = with a corresponding

$I$  = indicant of cohesion for each pair.

$_{.P}$  = The same population is classified into plurels in 2 ways or systems

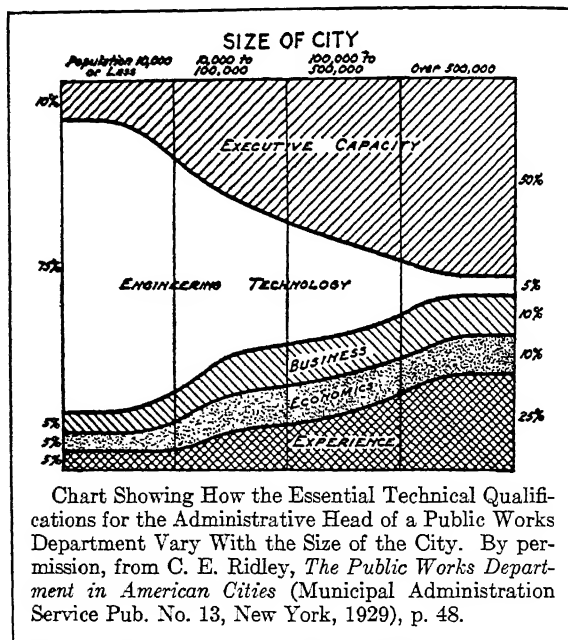
Comment:

This matrix is a mutual and, therefore, a symmetric one. The cell entries are the number of similar votes on bills by members of a legislature. The study develops a technic of analyzing legislative blocs, one form of co-acting sociological groups.

The situation is on the borderline of those defined in S-theory as "groups" by the quantic of P<sup>2</sup>, since the legislators are not stimulating and responding to each other directly (as in interperson attitudes), but rather their common and

mutually reinforcing responses to a common stimulus situation are measured in the cell entries. Therefore, the term "co-acting" rather than "interacting" is used above to describe this behavior. (Compare Eubank, p. 306, "Co-action," Ref. 25.)

## S. 11



Ref.: Pfiffner, John M., *Public Administration*, The Ronald Press Co., N. Y., 1935, p. 104.

Descriptive formula:  $S_{11} = {}^1P :: {}_pP : {}_iI_1$

Quantic number = 0;1;0;2

Legend:

$S_{11}$  = The situation  
records

$P$  = specified sizes of population  
with corresponding

${}^1P$  = a Public Works Director  
of

${}_iI_1$  = % of qualifications  
of each of

$|_p|$  = 4 classes of cities  
of

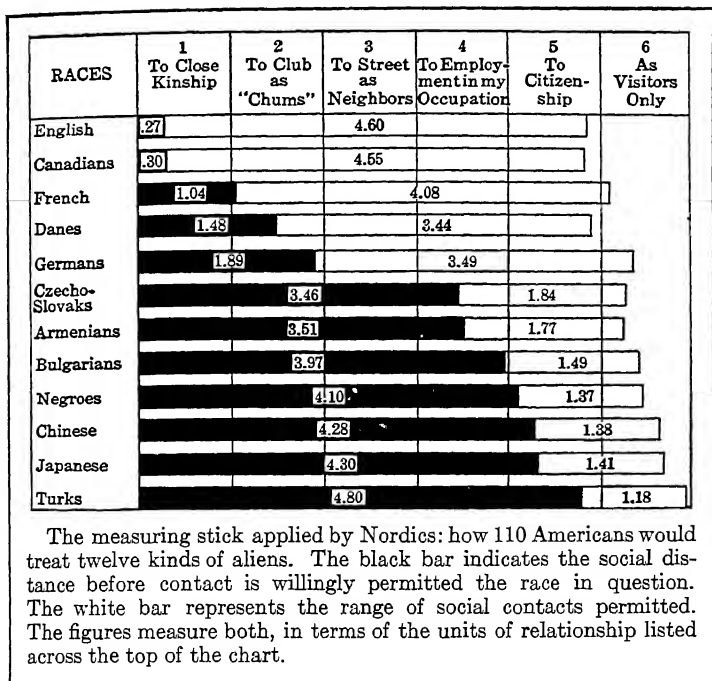
$|_i|$  = 5 kinds

Comment on notation:

This is an interrelation matrix of the person-plural type. Its one array of one person classified against 4 city plurals is subdivided by 5 kinds of qualifications making a third-degree matrix of order  $1 \times 4 \times 5$ . This is symbolized by the three descripts,  $|_1| \times |_p| \times |_i|$ .

Note that the  ${}^1P$  asserts a particular but unidentified person, i.e., a type. The officer, not the personal name of the officer, is recorded.

## S. 12



Ref.: Bogardus, E. S., "Social Distance—A Measuring Stick," Survey, Vol. LVI, No. 3, May 1, 1926, p. 170.

Descriptive formula:  $S_{12} = P, : \underline{P}_q : (I)_1$

Quantic number = 0;1;0;11

Legend:

$S_{12}$  = The situation

(I) = an index of social distance

is a record of

of each of

$P$  = a plurel of 110 Americans

$|_1$  = 2 types

who have, towards each of

$\underline{P}_q$  = 12 alien plurels (of unstated size)

*Comment on notation:*

This is a third-degree matrix of order  $1 \times 12 \times 2$ , as denoted by the descripts  $|, \times |_q \times |_i$ . It is a bipartite matrix of a single array  $|$ , i.e., a matrix of one-way interrelations of one plurel towards 12 other plurels. The black bars reflect a steeply sloping surface which with only one array becomes a line. The white bars show a surface which is nearer a level, or more of the equalitarian type.

The situation has 15 dimensions as shown by the sum of the aggregative descripts,  $|, + |_q + |_i = 15$ . See n of Eq. 52, Ch. II.

Note that the relation is one of contact not interrelation and is so symbolized by the single colon between the P's (instead of the double colon) and by the quantic digits of 11 instead of 2. The one-way attitudes of the Americans towards aliens are recorded without the reverse attitudes of those aliens towards Americans. If the two-way relations had been recorded here, the interrelations would be fully developed and the two digit 1's of the populational sector would be added to write a 2 as the quantic digit as then the populational product  $P^{+1} \times P^{+1} = P^{+2}$  would be fully developed.

## S. 13

Characteristics	Frequency with Which Characteristics Were Attributed to		
	Student A by 28 Fellow- Students	Student B by 20 Fellow- Students	25 Different Students by 51 Fellow-Students*
Sympathetic	15	11	44
Dignified	13	6	37
Friendly	12	14	46
Fair	10	10	36
Initiative	10	6	43
Intelligent	8	5	36
Social minded	8	2	22
Self-confident	8	6	39
Sincere	7	7	33
Dependable	6	2	22
Witty	6	0	18
Good speaker	5	4	12
Tactful	5	2	21
Vital	3	4	41
Good sport	3	3	15
Courageous	2	0	10
Aggressive	0	2	14
Optimistic	0	5	27
Honest	0	0	19
Attractive appearance	0	0	17
Positive	0	1	10
Democratic	1	0	9

\* NOTE: The figures in this column were obtained by summarizing the replies of fifty-two students. These replies were selected by a process of random sampling from about two hundred which were the most complete and the most intelligently given. Twenty-five different students were ranked either once or more than once by this group of fifty-two students, but no student was ranked more than three times.

*Descriptive formula:*  $S_{13} = P_p : P : {}^1\textcircled{I}_i$

*Quantic number* = 0;1;0;11

*Legend:*

$S_{13}$  = The situation

records

$P$  = population  $\left\{ \begin{array}{l} 1, "A" \\ 1, "B" \\ \text{and 25 re-} \\ \text{spectively} \end{array} \right.$

$P_p = 3$  student plurals  $\left\{ \begin{array}{l} 28 \\ 20 \\ 51 \end{array} \right.$

${}^1\textcircled{I}$  = all-or-none attributes

$|_i$  = of 22 kinds

who express, each towards a  
corresponding

*Comment:*

In analyzing one-way interrelation matrices such as this, the active party is normally written first, as the subject, the passive or recipient party is written next, and the attitude or connecting relation follows as dependent on the previous descriptis.

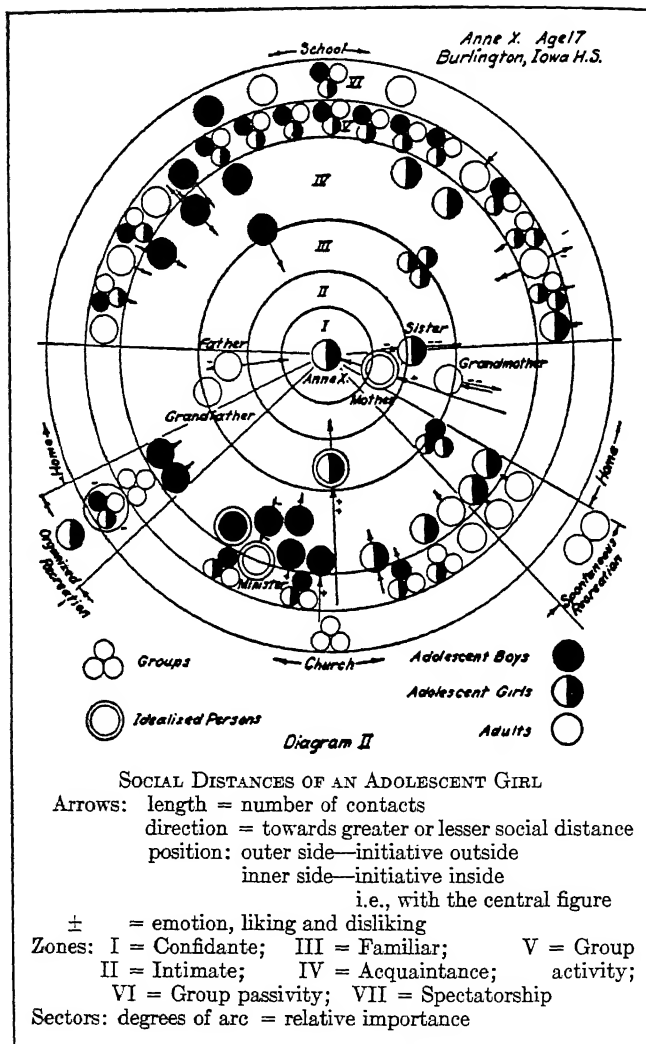
*Comment on notation:*

This matrix is a fragmentary one. It is a selection of the three main diagonal cells of a  $3 \times 3$  bipartite matrix of person-plural and plural-plural interrelations as follows:

	<i>Student A, 'P</i>	<i>Student B, ''P</i>	<i>25 Students, P<sub>v</sub></i>	} <i>I,</i>
$P_p = 28$ fellow-students	15			
$P_{''} = 20$ fellow-students		11		
$P_{'''} = 51$ fellow-students			44	

= the particular characteristic labeled "sympathetic"  
(and repeated for 21 other characteristics)

## S. 14



Ref.: Runner, Jessie R., "Social Distance in Adolescent Relations," *American Journal of Sociology*, Nov., 1937, Vol. XLIII, No. 3, p. 436.

*Descriptive formula:*  $S_{14} = 'P :: {}^pP_{|_p} : I_i$

*Quantic number* = 0;1;0;2

*Legend:*

$S_{14}$  = The situation

$I$  = indices of relationship

records

'P = for a particular person

:: = cross-classified

with

$|_i$  = of 5 kinds

${}^pP$  = 107 other persons

$|_p$  = in 6 plurels (characterized by sex, adulthood, idealization, or unspecified grouping, and institutional type)

{ number of  
contacts  
direction of  
contacts  
initiative  
location  
emotion  
intimacy

## V. NOTES

1. Further interrelations may be between groups, as in a social distance test (see S. 12, Ch. II and S. 1, Ch. VII) symbolized by:

$P_p :: P_P : (I) =$  formula for interrelations of groups (Eq. 2, Ch. VII)

In this case the number of cells in the matrix is the product of the plurel scripts which is  $p^2$  in the complete square matrix of interrelations between plurels.

Again, a large field of interrelations is that of the relations,  $i$  in number, between a person and a group (whether he be a member of that group or not) in:

${}^pP :: P_P : (I)_i =$  interrelation of parties (persons and plurels) interrelated by  $|_i$  characteristics (Eq. 3, Ch. VII)

Finally the *whole* matrix would combine *all* the interrelations of *all* persons and *all* groups in the specified population.

${}^pP_p :: {}^pP_P : (I)_i =$  complete formula for interrelations (Eq. 4, Ch. VII)

This becomes a formula for the organization of society when  $|_i$  represents a complete catalogue of all kinds of relationships, and when the population scripts represent all possible pairs of persons, pairs of groups, and person-group pairs. The symbolizing of societal structure is simple; the difficult scientific task is (a) devising objective indices for a greater proportion of the myriad kinds of interrelations that enmesh us, than have yet been devised, and (b) applying them more completely towards the goal of all possible pairs of parties in the population studied.

2. The interrelation of two persons is a product and classified in S-theory as a multiplication of  $P \times P$ . The justification for this is more psychological than mathematical at present. Psychologically, it is axiomatic that an interrelation depends on factors in both parties to the relation. Mr. X's liking for Miss Y depends on factors in him and in her, the *product* of which is the interrelation between them. The interrelation is observable in their behavior and may be recordable in the indices in their two cells of the matrix, but the factors in the two of them which created this product are often less easily observed and

have to be inferred. The matrix form may prove to facilitate accurate inferences as to such interpersonal factors, just as the matrix of intercorrelations is facilitating analyses into components.

Mathematically the interrelations can be summarized in an index which is a variant of the coefficient of contingency involving squares, and hence a quantic of 2. But in the absence, as yet, of a rigorous mathematical justification, any cross-classification of indices in S-theory is defined as a logical product of such quantitative or qualitative indices. (A matrix cross-classifying two independent indices is to be distinguished from a matrix subclassifying one index by another.) When the analysis of a recorded situation then yields a descriptive formula of the type:

$\{I\}_1 :: \{I\}_1 =$  formula for any cross-classification of indices (Eq. 5, Ch. VII)

its quantic formula is:

$$|s = (I)^2$$

It should be noted that the ordinary arithmetic product  $3 \times 2 = 6$  is a cross-classification of cardinal units when geometrically represented:

1			
2			
	1	2	3

The three equal linear units of one factor are subdivided into, or cross-classified against, the two equal linear units of the other factor with an area of six units as the product. Here the cross-classification is of cardinal units, but a cross-classification of ordinal units, or even of qualities ( $I^0$ ), can be considered as a more general type of product.

Geometrically, a product implies an area. The correlation coefficient or scalar product was seen to imply the area of a parallelogram (see S. 13, Ch. VI.) A similar parallelogram can readily be constructed for the interrelations in a population when some percentage type of index of the interrelations is graphed as a cosine of an angle, as in the example S. 12, Ch. II, where the angle is  $40^\circ$ . (See next note in this chapter.)

3. This has been done, for example, in S. 12, Ch. II, where the social distances between religious groups in Syria were expressed on a percentage scale, on which 0% is the friendliest possible attitude, i.e., no social distance, and 100% is the most hostile attitude possible, i.e., maximal distance. The average inter-religious distance was 41%. Of course, this can be arbitrarily diagrammed for a geometric interpretation by taking the percent as a  $\sin^2$  and finding the corresponding angle.

$$\sqrt{.41} = .64 = \sin \theta \quad \theta = 40^\circ$$

As 41% is 41% of complete alienation, the coefficient of alienation,  $k$ , which is the sine of an angle, is taken. This is the complement of the correlation and



As usual, the first  $\mathbb{P}_p$  denotes the row plurels,  $|_p$  in number, each with a certain number of persons,  $P$ , where  $P$  is a variable having  $|_p$  different values. The second  $\mathbb{P}_P$  denotes that, corresponding to each row plurel, there are column plurels,  $|_P$  in number, each with  $P$  persons in a cell where this second  $P$  is a variable taking  $|_p \times |_P (= p^2)$  values in all. The double colon shows the correspondence to be reversible in that it makes no difference whether each row is subdivided into columns, or each column into rows.

9. Many other variants of leadership phenomena can be quantitatively determined from the matrix. Thus, any ridge can be defined by the two equations of a line in a three-dimensional space, whose co-ordinates are  $P$ ,  $P'$ , and  $(I)$ . When the ridge is rectilinear this simplifies to:

$$P' = 'P \quad (\text{Eq. 9a, Ch. VII})$$

$$(I) = aP + '(I) \quad (\text{Eq. 9b, Ch. VII})$$

Eq. 9a defines a vertical plane in the array of the particular person, ' $P$ . In this plane, Eq. 9b defines a straight line whose ordinate is  $(I)$ , whose abscissa is  $P$ , whose slope is  $a$ , and whose intercept on the  $(I)$  ordinate is ' $(I)$ . Such a best fitting line can be used to describe a leader, or, for that matter, to describe any person in the interrelation matrix. Furthermore, by adopting some type line as an hypothesis, the goodness-of-fit of any person to that type can be measured by the appropriately calculated probability coefficient. This may prove a fruitful field for research.

Of course, the above equations for a ridge apply, with suitable shift of notation, to either a matrix of interrelated persons, or of interrelated plurels, i.e., to any group defined by some interrelation matrix. Note should be taken of the fact that, since the parties of the matrix are usually commutative, some definite sequence of rows and of columns must be fixed (either arbitrarily or by some criterion), or else the best fitting line will not be unique.

10. See Chapter I of Ref. 77 for a simple exposition of matrix algebra for psychologists and social scientists.

11. Similarly, when the row and column parties are plurels instead of persons, the diagonal matrix becomes:

$$\mathbb{P}_p : \mathbb{P}_P : (I) = (I)_{p:P'} \text{ plurels in isolation} \quad (\text{Eq. 10b, Ch. VII})$$

The left-hand member of Eq. 10a is a matrix in rectangular notation; the right-hand member is a matrix in "brief" S-notation. The left member of Eq. 10b is in "full" S-notation, while the right member again is in "brief" S-notation. "Brief" S-notation is fully explained in Chapter IX. It uses cross scripts, i.e., attaches the scripts of one sector ( $p$  here) to the base index of another sector ( $I$  here) for economy of symbolizing. The symbol  $p:P'(I)$  states that there is an index for each of  $p|$  persons towards himself. If the relation were an index for each of  $p|$  persons towards some other person among those  $p|$  persons, the symbol would be  $p:P'(I)$ ; while if it meant relations towards some person in a different population (such as employees each to his employer, or husbands each to his wife) the symbol would be  $p:Q'(I)$ .

The brief formula for

Eq. 1, Ch. VII, is  $p :: p(I)$

Eq. 2, Ch. VII, is  $(I)_{p :: p}$

Eq. 3, Ch. VII, is  $p : (I) : p$

Eq. 4, Ch. VII, is  $p :: p(I_i)_{p :: p}$

12. This formula, written in symbols, is:

$$100^{M(p :: (P-P')I)/M(p : P'I)} = (I)_I = \text{an isolation index} \quad (\text{Eq. 11, Ch. VII})$$

The formula as written assumes persons as the interrelated parties. Plurels, of course, may be the interrelated parties by shifting the person scripts to plurel scripts, as in Eq. 10a to Eq. 10b.

13. Among the uses of matrix algebra may be noted such processes as, pre-multiplying and postmultiplying matrices from one population, but involving different indices, in order to analyze, or isolate uncorrelated interrelations as components; the determination of array vectors in the matrix, and their scalar products and spatial relations leading to further insights; the reduction of the rank of the matrix resulting in more parsimonious re-expressing of the interrelations in terms of fewer and more elementary interrelations.

The mathematical aspects of the interrelation matrices are given more space here than the sociological, because the mathematical aspects are less familiar to the sociological reader to whom this book is addressed. It would be unnecessary to elaborate for him the various types of societal isolation, but their manipulation in diagonal matrices needs fuller exposition. This point of view prevails throughout the present treatise on quantitative systematic Sociology. It is assumed that most sociologists know sociological theory better than statistical theory, and that in weaving them together, therefore, the latter needs the emphasis.

14. As before when the interrelations are between groups, the formula becomes:

$$(I)_{p : (P-P')} \quad (\text{Eq. 12b, Ch. VII})$$

The single colon, denoting one-way dependence, distinguishes contact from interrelation, denoted by the double colon of two-way dependence. The brief S-formulae on the right of Eq. 12 state that there is an index of relation between each of the  $p$  persons (or the  $p$  plurels) and each of the others, excepting himself.  $(P-P')$  or  $(P-P',)$

It should be noted that the contact matrix is non-commutative, i.e., the sequence of the arrays is important. The isolation and the interaction matrices are commutative, that is, it makes no difference if the sequence of columns is changed, or if the sequence of rows is changed. But for contact, the first party in the first row can make contacts with, i.e., influence, all the remaining  $p - 1$  parties, the second party listed in the second row can only make contacts with the remaining  $p - 2$  parties, for if the second party contacts the first there will then be two-way contacts constituting interaction between them. Similarly, the third party in row 3 will contact the remaining  $p - 3$  parties only; and so on, until the last party contacts no one, but is contacted by all. In practice,

when studying contact phenomena, only certain cells or arrays are usually studied; but if completeness is desired, the sequence of listing the arrays becomes important. Usually a rectangle of cells, selected from the triangular area which lies on one side of the main diagonal, defines  $p$  parties who contact  $q$  other parties.

15. All these numbers of cells in these triangles, main diagonals, total matrix, etc., are but special cases of the general formula for the number of combinations of  $p$  things taken two at a time,  $p!/2(p-2)!$ , or the number of their permutations,  $p!/(p-2)!$ . The number of the permutations is twice the number of combinations, since the data are taken in pairs and each pair has two permutations.

16. This index is:

$$100M(p:(P-P')I)/M(p:(P-P')I)' = \text{an index of contact} \quad (\text{Eq. 13, Ch. VII})$$

It is possible to combine Eqs. 11 and 13 into one percentage which is zero in isolation, rises to 50% in pure contact, and reaches 100% in symmetric interrelation. But such an index requires two formulae—one for the 0%-49% range, and another for the 50%-100% range, because contact is not on the rectilinear continuum defined by the two extreme points of isolation and symmetry. Contact is an asymmetric situation. It is entirely possible for isolation to shift to a high degree of symmetric interrelation without passing through the purely one-way set of relations which define contact.

17. The relation between interrelation matrices and a contingency table might be noted, as the two tend to be confused in the case of bipartite populational interrelation matrices. Here the analyst is sometimes puzzled whether to write the descriptive formula as:

$$P_p :: P_q \quad \text{the bipartite populational interrelation matrix,} \\ |^s = 0;0;0;2 \quad (a)$$

$$\text{or} \quad I^q_P :: I^q_P : P \quad \text{the contingency table, } |^s = 0;0;0;1 \quad (b)$$

On writing explicitly the implicit indicial-population products in (a) it becomes:

$$I^q_P :: I^q_P \quad |^s = 0;0;0;2 \quad (c)$$

which is close to (b) in meaning, but with a different quantic. The quantic of (b) classifies the situation under plurels,  $P^{+1}$ ; the quantic of (a) and (c) classifies the situation under groups,  $P^{+2}$  (interrelated persons). Which is the more correct description will not be decided by the formula but only by the judgment of the sociologist writing the formula. He must apply the definitions of a plurel as persons with a common characteristic, and of a group, as persons in psychic interrelation, i.e., in a stimulus-response relation to each other, conscious of each other, interacting, having a "we-feeling," etc. Clearer criteria for borderline situations may be needed.

To fix thinking, consider a concrete example of the abstractions symbolized above. Let one student population be classified into plurels on the basis of each of three characteristics, namely: eye color (in 3 classes), political party (in 4 classes), and fraternity membership (in one of 5 fraternities). If now eye color

is cross-classified with political party, and the number of students characterized by a particular eye color and a particular party is entered in each cell of the matrix, a contingency table is obviously being made to determine the relationship between two qualitative characteristics, rather than to determine the psychic interrelations of groups. (b) is the appropriate formula and the situation belongs under plurels  $P^{+1}$ , in the quantic classification. Next, cross-classify the political parties with the fraternities. Either (b) or (c) describes this matrix, for obviously a contingency coefficient can be calculated, and yet it can be argued that the significance of this situation is not the degree of contingency between the two characteristics, "party" and "fraternity," but the psychic interaction, indicated by overlapping membership, between political groups and fraternity groups. Here the cross-classified plurels are composed of persons who are conscious of each other and of their group-unity, who are interacting, who obviously satisfy the sociological criteria of a group and transcend a mere plurel. If the analyst judges that the psychic interrelations of whole persons and groups to each other is the point of the data as presented, he should describe it by formula (a); if he judges that correlation of the qualitative characteristics of plurels is the point, he should describe the situation by writing formula (b).

Borderline, or ambiguous situations, may be expected, especially where the data are inadequately captioned, causing disagreement between analysts in classifying the situation. The function of the symbols of S-theory here is largely to force the issue into the open by requiring exact classification, and revealing any ambiguities as problems for further research in defining classes.

For some examples of situations on the borderline of plurels,  $P^{+1}$ , and groups,  $P^{+2}$ , the student might note the following on which the first analyses by Mr. Lundberg, the author, and his graduate student assistant, disagreed, and which had to be discussed to reach a consensus as to its quantic classification (as presented here):

*S. 2, Ch. X. Reorganizing army departments*

$|^s = 9;0;0;1$

Instead of a simple change in an aggregation of plurels, could the lines connecting the old to the new departments be interpreted as an all-or-none indicant of interrelation, such as "inclusion" and "non-inclusion," so that the situation should be described by a bipartite interrelation matrix of every old Department, cross-classified with every new Department, and "included" or "not-included,"  $^{1,0}I$ , be the cell entries? Parsimony favored the interpretation of changing plurels.

Note that the ambiguity between plurels and groups occurs in these cases and in organization charts where the exact nature of the relations of the parts is not clearly expressed in the data (as presented in print for the analyst to classify).

*S. 3, Ch. IV. Social security administrative departments*

$|^s = 0;0;0;1$

Instead of a subclassification of plurels, can this situation be interpreted as Government Departments interrelated with client plurels by some sort of "responsibility" relation? Since, in the situation as recorded, the lines connecting the Departments and clients do not even have the nature of the sug-

gested relation named, let alone measured, the simpler interpretation of subclassified plurels was adopted on grounds of parsimony. The analyst should avoid reading into the situation more than is recorded in it as printed.

S. 7, Ch. IV. *An industrial organization chart* |<sup>s</sup> = 0;0;1

Instead of a population subclassified into a hierarchy of plurels, could this situation be interpreted as a cross-classification of every economic plurel with every other one? The indefiniteness of the attribute, or index of interrelation, to go in each cell of such an interrelation matrix led to the rejection of this interpretation in favor of the simpler one of subclassified plurels.

18. The general equation of such a plane is:

$$A^{\Sigma P \Sigma P} + B^{\Sigma P \Sigma P} + C(I) + D = 0 \text{ the monoplanar interrelation surface} \\ (\text{Eq. 15, Ch. VII})$$

where  $\Sigma P \Sigma P$  (the number of actor persons and plurels), is the row co-ordinate; where  $\Sigma P \Sigma P$  (the number of recipient persons and plurels), is the column co-ordinate; where (I) is the third altitude co-ordinate; where A, B, and C are coefficients of slope, and where D is a constant determined by the intercepts on the co-ordinates.

19. A formula for the equalitarian matrix is:

$$\sigma(P :: P I_P :: P) \doteq 0 \text{ the equalitarian, or level, interrelation surface} \\ (\text{Eq. 16, Ch. VII})$$

Formulae indicating the singly sloping surface are:

$$\sigma(P \Sigma I) \doteq 0 \quad (\text{Eq. 17a, Ch. VII}) \\ \sigma(P \Sigma I) > 0 \quad (\text{Eq. 17b, Ch. VII})$$

where  $\Sigma$  as usual denotes the aggregation of row totals and  $P$  the aggregation of column totals. For the doubly sloping surface, Eq. 17a would have the "greater than" sign ( $>$ ) replacing the  $\doteq$ .

20. At this point, having proposed an operational definition of "community," proposers of verbalistic definitions might be challenged to an experimental test. Suppose that under an impartial jury of sociologists, assisted by a grant-in-aid, and graduate students equally well trained in S-theory and in alternative theories defining a "community" (by the aid of fellowships for this purpose if necessary), a sample, say a 100, of descriptions of communities and non-community plurels were made up, using criteria previously agreed upon by the authors of the rival theories.

The trained graduate students would then classify these 100 descriptions as being communities or not-communities, or into more refined classes of kind or degree of community. This classifying would be done independently by each student, once on the basis of the S-theory definition, and again on the basis of any specified alternative definition. Percentages of agreement between judges, or other indices of consistency of classification would then be computed and compared for the two definitions. Which type of definition, operational or "synonymistic," would be experimentally demonstrated to be most reliable?

This would test the reliability, but not the validity or apparent sociological

significance, of the two definitions. But what is the use of any sociological or any scientific concept, however significant it may seem to our current thought-ways, if it cannot be reliably determined, and no technics are in sight for improving such reliability? All technics, for increasing the reliability of the use of verbalistic subjective concepts, seem to the author to tend in the direction of objectivity and operational concepts. The very definition of reliability as agreement on repetition of the observations involves an operation! Unreliable concepts should be discarded and a science built on concepts that are reliably determinable, i.e., on *facts* not on *opinions*, whatever their apparent lack of significance at present.

21. If two parties ( $p = 2$ ) have three possible interrelating attitudes ( $i = 3$ , namely: (a) "attraction," (b) "indifference," and (c) "repulsion") the structures of this population may be represented as follows: letting each small letter (a, b, c) denote an attitude and each capital letter denote one structure in a pair of persons:

"Pair structures"	Attitudes in a pair	"Pair structures"	Attitudes in a pair	"Pair structures"	Attitudes in a pair
A	a a	D	a b	D'	b a
B	b b	E	a c	E'	c a
C	c c	F	b c	F'	c b

Here in the first three "pair-structures," A, B, and C, the attitudes are mutual between the two members of the pair. The remaining six "pair-structures" D-F' have three different combinations (D, E, and F) of attitudes but six permutations (D, E, F, D', E', F'). There are altogether nine permutations (A-F') of the three attitudes when these are taken two at a time, allowing repetitions, but excluding intra-relations.

The number of possible combinations is determined by the number,  $r$ , of possible pairs of parties and by the number,  $n$ , of "pair-structures" such as A-F above. Thus if there are three parties ( $p = 3$ ), there are three possible pairs ( $r = 3$ ); and if there are but two attitudes (a and b above) yielding three "pair-structures," A, B, and D; then the following ten "group-structures" are possible:

A A A	A A B	A B D
B B B	A A D	A D D
D D D	A B B	B B D
		B D D

The generalized formula for the number of possible "group-structures" is the number  $Q_3$  of combinations of  $n$  things taken  $r$  at a time, allowing repetitions:

$$Q_3 = \frac{(r + n - 1)!}{r!(n - 1)!} \quad (\text{Eq. 19a, Ch. 7})$$

To calculate this,  $r$ , the number of possible pairs of the  $p$  parties, is the number of combinations of  $p$  things taken two at a time:

$$r = \frac{p^2 - p}{2} \quad (\text{Eq. 19b, Ch. 7})$$

This is the number of cells in one triangle above the main diagonal in the inter-relation matrix. It is also the number of different terms in the square of the multinomial of  $p$  terms.

The number,  $n$ , of possible "pair-structures" is the number of combinations of  $i$  things taken two at a time when repetition is allowed:

$$n = \frac{i^2 + i}{2} \quad (\text{Eq. 19c, Ch. 7})$$

Thus for the case above of three parties with two possible attitudes,  $p = 3$ ,  $i = 2$ , and  $r = 3$ ,  $n = 3$  and  $Q_3$  is:

$$Q_3 = \frac{(3 + 3 - 1)!}{3!(3 - 1)!} = \frac{5!}{3! \cdot 2!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = 10$$

which are the ten "group-structures" represented by AAA,  $\dots$ , BDD above.

Some values of Eq. 19a, are:

		i				
		1	2	3	4	5
p	1	$Q_3 = 1$	$Q_3 = 2$	$Q_3 = 3$	$Q_3 = 4$	$Q_3 = 5$
	2	1	3	6	10	15
	3	1	10	56	220	680
	4	1	28	462	5,005	38,760
	5	1	66	3,003	92,378	1,961,256
				$2 \cdot 1$	$p! \quad 3!$	$p! \quad 4!$

The first row (of one plurel) expresses self-attitudes only and becomes meaningful when  $P$  represents plurels instead of persons and these first row entries would then be in-group attitudes. The first column has only one attitude (such as "will associate with") without even its opposite, so that here there is no break-down whatever of the total population into subplurels).

22. The formula is:

$$Q_i = i^p = \text{number of permutations of } p \text{ parties having } i \text{ possible interrelations (including intra-relations) (Eq. 20, Ch. VII)}$$

Some illustrative values are:

		i			
		1	2	3	4
P	1	$Q_1 = 1$	$Q_1 = 2$	$Q_1 = 3$	$Q_1 = 4$
	2	1	$2^4 = 16$	$3^4 = 81$	$4^4 = 256$
	3	1	$2^9$	$3^9$	$4^9$
	4	1	$2^{16}$	$3^{16}$	$4^{16}$

A simple derivation of this formula is from our fundamental interrelation matrix where the  $i$  in each cell represents either the  $i$  possible degrees, or  $i$  possible kinds, of relation of the row party to the column party.

	$P_1$	$P_2$	$P_3$	----	$P_{p'}$
$P_1$	(i)	i	i	i	i
$P_2$	i	(i)	i	i	i
$P_{i_3}$	i	i	(i)	i	i
$\vdots$	i	i	i	(i)	i
$P_{p'}$	i	i	i	i	(i)

$= \underline{P}_p :: \underline{P}_p : \underline{I}_i$  (Eq. 21, Ch. VII)

Since the  $i$  possible attitudes of party  $P_1$  to party  $P_2$  can coexist with any of  $i$  reciprocating attitudes this yields  $i \times i$  or  $i^2$  possibilities. But each of these  $i^2$  possibilities can coexist with any of the  $i$  attitudes of  $P_1$  to  $P_3$  making  $i^3$  possibilities; and all these can coexist with  $i$  possible attitudes of  $P_3$  to  $P_1$  making  $i^4$  possibilities. Continuing thus,  $i$  is taken as a factor once for each of the cells,  $p^2$  in number, which means that  $i$  is raised to the  $p$  squared power (Eq. 20).

If intra-relations are excluded the  $p$  main diagonal cells are eliminated, making:

$Q_2 = i^{p(p-1)} =$  number of permutations of  $p$  parties having  $i$  interrelations (when intra-relations are excluded) (Eq. 20a, Ch. VII)



## *PART IV*

### THE SPACE SECTOR, $\mathbb{L}_1^1$

*studying situations defined by  $S = T^0$ ;  $I^i$ ;  $L^{l \neq 0}$ ;  $P^p$*



## Chapter VIII

### DENSITIES, $L^1$

#### I. PHYSICAL SPACE AS A SOCIOLOGICAL CATEGORY— INDICES OF LENGTH, $L$

In the early stages of development of the S-theory, only the components T, I, and P were used, and physical length was classified as one subtype of indicant, a spatial characteristic of the environment of people. As the analysis of maps and other recorded situations involving areas proceeded, it became evident that to treat physical space as an indicant was unsatisfactory, and that it should be symbolized separately as a sector, i.e., a type of index, co-ordinate with time, population, and characteristics. For in the first place, space occurred as an index in societal situations with a far greater frequency than any one other indicator. In the second place, it usually occurred as an area denoted by an exponent of two,  $L^2$ , whereas the other indicants of the same situation usually occurred with an exponent of one. This heterogeneity of powers of the indicatory indices in one situation proved confusing.

In the third place, the scalar product had been adopted as the more general form of multiplicative compounding of societal indices, since the scalar product denotes correlation, and any societal phenomenon is correlated, even if in zero amount, with any other societal phenomenon. But to multiply any two indicants of societal phenomena together into an ordinary arithmetic product usually produces a meaningless result. The scalar product, which is geometrically represented by the area of a parallelogram, has as its limiting case, when correlation is perfect ( $r = 1.00$ ), the arithmetic product, which is geometrically represented by the area of a rectangle. Societal indicants are almost never perfectly correlated with each other. Even the reobservation of one indicant usually shows a reliability correlation of less than unity. But in physical area the two dimensions of length and breadth are conventionally taken as perpendicular to each

other, and hence are defined by two collinear vectors representing perfectly correlated variables. (Note that in a scalar product the parallelogram and rectangle are defined by one vector and the *normal* to the other vector. Study S. 13, Ch. VI for this bit of geometry.) Spatial dimensions in thus showing perfect correlation with each other were in marked contrast to all other indicants which showed all degrees of imperfect intercorrelations.

For indicants other than length, a quantic of 2 denotes that the situation has been observed and analyzed beyond the purely qualitative stage, beyond the next quantitative stage into the third stage of correlations between the characteristics. The indicatory exponent thus states the degree of operational penetration that has been achieved into the relationships and patterns of the data, whereas the exponent of 2 on a length component has a different meaning of "area," or relations in a geometric plane. By assigning to length a separate symbol and sector of S-space, correlation,  $I^2$ , and area,  $L^2$ , were clearly distinguished.

In addition to the above technical considerations, other broader considerations played a part in adopting space as a fourth sector in S-theory. The primacy of time and space as the milieu of human life was considered. The importance of human geography and ecology in the sociological literature was considered. One of the chief subdivisions of Sociology is into Rural and Urban Sociology, and this is a classification essentially based on density of population, in which area is one of the two factors. Leplay's formula analyzing societal phenomena into categories related to Place, Work, Folk ( $L$ ,  $IT^{-1}$ ,  $P$  in S-notation) is still a useful scheme (Ref. 66). The analysis by economists of production into the factors of Land, Labor, and Capital, operating in time ( $L^2$ ,  $P$ ,  $I$ , and  $T$  in S-notation) is another example of well-tried analyses of some segment of societal phenomena which use physical space as a basic category.

The adoption of space,  $L^1$ , as a fourth sector, i.e., as a fourth basic category, seemed desirable in analyzing the graphs and tables in keeping with our definition of Sociology as the study of "the characteristics common to all classes of social phenomena." As stated in Chapter II, the number of sectors chosen is a matter of utility; it can be more or fewer. Systems can be built up using  $T$ ,  $I$ , and  $P$  alone, or  $T$  and  $I$  alone, or even  $I$  alone,

with all the others as subsectors. A fifth sector for monetary indicators may be desired in Economics, or for value indicators in Ethics, and may at any time be so symbolized by substituting  $M$  for  $I_M$ , and  $V$  for  $I_V$ . Experience in analyzing fifteen hundred quantitatively recorded situations sampled from all the social sciences has led to the belief that the four categories of time, space, people, and all the residual characteristics are the most suitable:

- a. for marking off societal phenomena from the non-societal phenomena of the other sciences,
- b. for classifying the characteristics common to all classes of societal phenomena, and
- c. for using classes (i.e., categories, sectors,— $T$ ,  $I$ ,  $L$ , and  $P$ ) which are most objectively definable, that is, which are least likely to overlap in meaning or be confused with each other.

This is an application of the three canons for any scientific classification, namely, a single basis, total inclusion of the field, and mutual exclusion of the classes.

Although space was, therefore, accorded the status of a fourth sector, our analyses of  $S$ -situations showed it to be both a sector which occurred less frequently than time, indicators, and people, and one which had less sociological importance than the other three.<sup>1</sup>

## II. THE NOTATION OF THE SPACE SECTOR—THE SCRIPTS ON $L$

### A. The Exponent, $L^1$

In situations where physical space is not involved, the  $L$  is nul ( $L_0^0$ ). It has an exponent of zero and zero descripts, and hence is not written in the descriptive formula, but is written as 0 in the quantic number. (See any  $S$ -situation appended to Chapters III–VII.) Points in space which are named but not located in any line, area, or volume are denoted by the point script and a zero exponent.

Lines whether straight, curved, or broken, representing routes, boundaries, rivers, latitude and longitude, or other co-ordinates, altitude contours, radii of zones, or other geographic, or geometric,

linear magnitudes are symbolized by an exponent of 1 on  $L$  and appropriate descripts.

As usual the particular classifier script,  $|$ , specifying the quality, i.e., kind of line, is understood to be present if it is not written (in Eqs. 2b and 2c, as also in Eqs. 3c and 3d, and 4b and 4c below).

Areas and regions of a surface, such as square feet of floor space, acres, zones, districts, countries, are denoted by an exponent of +2. (The + is understood, if not written.)

Volumes, or three-dimensional situations as in cubic units; latitude, longitude, and altitude; street, avenue, and floor of a building; are denoted by an exponent of 3.

A negative exponent, as usual, denotes that the length component is a divisor, or is taken as the unit, as in cost per kilometer,  $IL^{-1}$ , people per square mile,  $PL^{-2}$ , or dollars per liter,  $IL^{-3}$ .<sup>\*2</sup>

### B. The Class Script, $L_1$

The "pure" index of length with zero descripts is of one kind only. Length is length no matter in what direction it is measured or what kind of thing is measured. It may be expressed in diverse units, but if the units are standardized, i.e., equal and interchangeable, any unit of one kind has a constant ratio to any other kind of unit of length.

Whenever the length (or area, or volume) is qualified, however, as being a length of *some sort of thing*, as in inches of stature, miles of railroad, arable acreage, cubic capacity of skulls, etc., we have an implicit product of length and the attribute, or qualitative characteristic, whose length is stated:

$$I^0; L_0^{+1} = L_1^{+1} \text{ length of something, the product of an attribute and pure length} \quad (\text{Eq. 5a, Ch. VIII})$$

(Compare Eq. 9, Ch. III and Eq. 4, Ch. IV)

This implicit product, of course, may be generalized to products of a hierarchy of attributes and spaces of any of the three dimensions:

$$I_{i;j;k}^0; L_0^1 = L_{l;m;n}^1 \quad (\text{Eq. 5b, Ch. VIII})$$

$$\text{quantic number} = 0;0;1;0 \quad (\text{Eq. 5c, Ch. VIII})$$

\* For Eqs. 1a-4c, Ch. VIII, see notes at end of chapter.

In the great majority of societal data space occurs as such an implicit product, for the lines, areas, etc., are always named, or qualified by some adjective, by some position on a map, or by other qualitative differentia. As soon as the quality describing the space becomes observable in quantitative degrees so that the attribute becomes an indicant, that indicant must be written explicitly and never as an implicit product.<sup>3</sup> (See S. 17, 18, 19, 20, 21, Ch. VIII.)

The number of vectorial dimensions in the space sector is stated, as usual, by the class script. Strictly these are attribute dimensions, but in implicit products they may be conveniently treated as vectorial dimensions in the space sector, which would otherwise have but one vectorial dimension for any given exponent dimension. The total number of dimensions in the space sector is the product of the exponent and the class script.

$|^1 \times |^1$  = number of spatial dimensions of both class and exponent types, i.e., vectors and their perpendiculars  
(Eq. 6, Ch. VIII)

### III. SOCIETAL DENSITIES, (I) $L^{-1}$

#### A. Classification of Densities

In societal data space is important only in its relation to human life. This relation is usually expressed in the form of an explicit or an implied ratio. Thus the density of population of a country is expressed as the number of persons per square mile,  $PL^{-2}$ ; the crowding in tenements is expressed as square feet of floor space per person,  $L^2P^{-1}$ , or as persons per room,  $PL^{-2}$ ; the cost of building a highway is expressible as the dollars per 100 miles of road,  $IL^{-1}$ . All these ratios of some index and a spatial index may be termed "densities."<sup>4</sup>

(I) $L^{-1}$  =  $D_n$  = a density, any index per unit of space  
(Eq. 7a, Ch. VIII)

These densities may be classified in three ways: (1) according to whether the spatial index is the numerator or denominator in the ratio; (2) according to the exponent,  $|^1$ , of the length index; and (3) according to the sector of the second index, (I), in the density ratio.

The more frequent convention is to express the densities of some characteristic, or of a population, in spatial units, i.e., as  $x$  I-units or persons per square mile. When, however, the spatial index is in the numerator, as in cubic feet per person, acres per capita, miles per hour, the ratio may be called an inverted density.

$(I)^{-1}L^1 = Dn^{-1}$  = an *inverted* density, space per unit of any index  
(Eq. 7b, Ch. VIII)

The terms "obverse" and "inverse" densities are sometimes useful in specifying Eq. 7a or Eq. 7b respectively. For some purposes the inverse density is the more convenient, or socially significant measure, as in thinking in terms of acres per farm, annual gallons of drinking water per capita, kilometers per hour of traveling speed, etc.

The density may be in terms of any of the three dimensions of space, as a *linear* density of points on a line, or the usual *areal* density of the man-land ratio, or a *cubic* density as of smoke particles per cubic centimeter of air. Obviously societal density is not limited to the meaning of density in Physics, where it is the ratio of mass to volume.

Societal densities may be classified further according to the nature of the second component, as:

$PL^{-1}$  = populational densities (Eq. 7c, Ch. VIII)

$IL^{-1}$  = indicatory densities (Eq. 7d, Ch. VIII)

$TL^{-1}$  = temporal densities (Eq. 7e, Ch. VIII)

$LL^{-1}$  = lineate (and/or areal) densities (Eq. 7f, Ch. VIII)

Each of these will be briefly considered in respect to their function of systematizing societal data.

## B. Populational Densities, $PL^{-1}$

### 1. RURAL AND URBAN SOCIOLOGY

The density of population in a region is an index of the type of culture of the population. In a hunting and collecting culture the population is necessarily scarce. It becomes denser as the population goes on to a pastoral, to an agricultural, to a commercial, and finally to an industrial culture. This series of productive-cultures can be considered an ordinal indicant of "degree of utilization of the environment,"  $^1I$ , ranging from least utiliza-

tion in pastoral nomadism up to most utilization in modern industrialism. The relation of this indicant of environmental utilization to population density can be expressed by fitting the curve:

$${}^iD_n = 2.5(3^i) = \text{relation of density and productive-culture} \\ (\text{Eq. 8a, Ch. VIII})$$

Here the *i* in the exponent is numerically identical to the case script, *i*, which expresses the rank of each productive-culture as in the column headed <sup>i</sup>| of the tabulation below. This formula is seen to give a fair approximation to the densities of the cultures as specified by Miss Semple (Ref. 61.)

<sup>i</sup>	Productive-Culture	<i>D<sub>n</sub></i> (Persons per Sq. Mile)		<i>D<sub>n</sub></i> <sup>-1</sup> (Hectares per Person)	
Rank (Case script)	<sup>i</sup> I	by Semple	by Eq. 8a	by Semple	by Eq. 8b
0	Nomadism as in Arabia	2.5	$2.5(3^0) = 2.5$	104	$.5(3^5) = 121.5$
1	Agricultural, primitive	5-15	$2.5(3^1) = 7.5$	17-52	$.5(3^4) = 40.5$
2	Agricultural, colonial	25	$2.5(3^2) = 22.5$	10.4	$.5(3^3) = 13.5$
3	Agricultural, central European	100	$2.5(3^3) = 67.5$	2.6	$.5(3^2) = 4.5$
4	Agricultural, southern European	200	$2.5(3^4) = 201.5$	1.3	$.5(3^1) = 1.5$
5	(Agricultural and industrial) Industrial	500-800	$2.5(3^5) = 604.5$	.5-.3	$.5(3^0) = .5$

Since the area available for each person may be a more logical view of density than the number of persons per area, the inverse densities in units of the metric system stated as the number of hectares per person,<sup>5</sup> *D<sub>n</sub>*<sup>-1</sup>, are given by the parallel formula:

$${}^iD_n^{-1} = .5(3^{5-i}) = \text{relation of inverse density and productive-culture} \\ (\text{Eq. 8b, Ch. VIII})^6$$

Density may be related both as cause and as effect of cultural

factors for, on the one hand the density cannot exceed the food supply which that type of culture maintains, and on the other hand specialization and other cultural phenomena may depend in part upon the number of people amassed together. The correlation of density and cultural type is imperfect. Topography and climate, recency of the pioneering stage, and other factors up to, and including, human legislation keep the correlation from being perfect. Nevertheless, the density of a population tells much about their probable culture. This correlation is sufficiently important to be depended upon as a major factor in societal control. States desiring a larger population for military or other reasons, with a given territory, know that increased density can be maintained only as increased industrialization accompanies it.

A second correlate of density which is important in Sociology is the distinction between rural and urban culture with its corresponding division for study into Rural Sociology and Urban Sociology. Perhaps the most dependable and universal index differentiating rural and urban in all periods of history, or in all parts of the world, is a density, implied or explicit. Rural culture correlates with low density of population and urban culture with higher density. As density changes, the culture tends to change correspondingly. This correlation would seem to be so high that it is here proposed that the terms rural and urban be defined by densities below and above some conventionalized boundary point. Thus the point adopted by the United States Census is centers with fewer or more than 2500 inhabitants. The international standard boundary is at 2000 inhabitants. Owing to the tendency of people to nucleate, i.e., to reside collected in centers and not evenly scattered over an area, this boundary, stated in terms of a population in a center with the area it inhabits left indefinite, is more easily determined, though theoretically less elegant than a true density.

Since populations tend to collect at centers or points on the map (called villages, towns, or cities), the area can be indefinite and the areal denominator of the density ratio can be treated roughly as a constant, enabling relative densities to be compared by comparing the numerators of the density ratios, i.e., by comparing absolute populations in centers.

This definition of the rural-urban boundary resulted in dividing

the population of the United States into 50% urban and 50% rural (Ref. 56) in 1910, and into 44% rural in 1930.

The boundary point might with research become a matter of consensus of scientific opinion as some point most suitable for all countries. Thus from Miss Semple's tabulation a reasonable boundary stated in the metric system might be one hectare per person; or it might be defined differentially as the density in a given country which divides a predominantly farming population from a predominantly non-farming one. This would make the derivation of the boundary depend on culture, but thereafter it would be more objectively and easily determined by measuring densities. Or the average density might be taken, and rural and urban become synonymous with subaverage density and super-average density; or again, more exactly, the phenomena may be classified as finely as desired by specified class-intervals of density. In comparing densities the areas should be comparable, for a small city in an agricultural country will show a density higher than that of the most industrialized nation for the whole national territory.

Rural and urban plurels defined by their density might be termed *plurels of density* in distinction to the *density of plurels* which are defined by other characteristics. In the latter category belong such ecological studies as, determining juvenile delinquency rates per area in various city zones, or the densities in a region of races, income classes, or any other plurels. The difference between plurels of density and the density of plurels is whether the density defines the plurel or the plurel defines its density. If in the operation of collecting and presenting the data the density is the independent variable and the plurel the dependent one, it is a plurel of density; while if the plurels are the independent variable for each of which its corresponding density is observed, it is a situation involving density of plurels.<sup>7 \*</sup>

## 2. CROWDS

Another class of sociological phenomena in the definition of which space is an essential component is that of crowds. Von Wiese and Eubank (Refs. 25 and 78) in their systems of Sociology divide all plurels into three kinds:

\* For Eqs. 9a-d, Ch. VIII, see notes at end of chapter.

- a. the human category (class)—defined by a common characteristic =  $P$ , (or  $I^0P_0$ ) (Eq. 10a, Ch. VIII)
- b. the human aggregation, crowds, mobs, audiences—defined by spatial proximity =  $L^3 : P$  (Eq. 10b, Ch. VIII)
- c. the human interactivity or group—defined by psychic interaction =  ${}^pP :: {}^pP : (I)$  (Eq. 10c, Ch. VIII)

The first is the "plurel" discussed in Chapter IV for qualitative characteristics, and in Chapter V for quantitative characteristics. The third is the "group" discussed in Chapter VII for its static aspects. (For the dynamic aspect of all of these see Chapters X and XI.) The second is a "crowd," discussed below. Although the three kinds of plurels are discussed in separate chapters, it should be realized that they form a contiguous region of the quantic solid, S. 33, Ch. II. Plurels constitute the middle stratum of that solid, groups constitute the top stratum, and crowds are in the arrays where these strata intersect with the  $L^2$  and  $L^3$  planes (not diagramed in S. 33, Ch. II).

A crowd may be defined as any plurel in a space small enough for the persons to see and hear each other. (Radio amplifiers and television may, of course, extend the boundaries of a crowd at least in the one-way direction of contact, if not in the two-way direction of interaction.) Crowds may be classified on the basis of various indicators of emotion, or indicators of causes bringing people together, or indicators of the collective action resulting. Among other classifications the following one, based on space, belongs in this chapter.

Crowds may be classified according to the spatial location of the stimulus to which the persons are responding at the moment. From this point of view crowds are, (a) scattered, (b) nucleated, or (c) unified. When every person is going his own way, "attending to his own business," as in an ordinary throng of people on the street, they (the people) are a "scattered" crowd. Their foci of attention are scattered, their eyes are on all sorts of things, they are attending to each other or items in the environment in as many points of space as there are persons in the crowd. The next class is when they form "knots," or minor gatherings, here and there, in each of which several persons are attending to the same stimulus whether a shop window, a tram doorway, a soap-

box speaker, or a circle of friends in conversation. Here the crowd is nucleated, or collected together into as many nuclei as there are points of stimulation to which people are attending. The third class is when all the persons within eyesight or hearing are attending to one stimulus at one point in space. A crowd is "unified" when every one is listening to one speaker, watching one movie screen, fleeing from one mad dog, marching past a reviewing stand, etc. This classification can be made by the aid of a movie film of a crowd showing what each person is looking at, since, in general, our gaze indicates the spatially located stimulus in the immediate environ to which we are at least partially responding. This behavioristic index of the degree of unity in a crowd is useful for prediction of their behavior. Proportionately as a crowd becomes unified, as defined above, the persons in it are likely to behave alike and simultaneously, "all in one piece," as a collective unit.<sup>8 \*</sup>

For an example of such a unified crowd and an example of photographic methodology in classifying crowds, study S. 25, Ch. VIII. A street crowd during a revolution is being dispersed by a machine gun on the roof of a building at the corner of intersecting streets. As the crowd becomes aware of the location of the gun and that it is mowing people down, they are shown in the process of becoming unified in the common reaction of flight from the single death-dealing stimulus on that roof.

### 3. A CLASSIFICATION OF REGIONAL PLURELS BY SIZE

In addition to crowds, another classification of plurels involving space, which have been called human aggregations in Eq. 10b, Ch. 8, is that of regional plurels. As these regional plurels, (which are usually also regional communities, as defined in chapter 7) illustrate a use of S-notation in reducing an ordinal series to a neat formula, it will be sketched here.

Regional plurels may be classified by the size of population—which is correlated with the size of area. Thus the classes of village, town, city, nation are common terms but are of variable meaning as to the size of population involved. The chief standardized boundary is in the census distinction of "rural" and "urban,"

\* For Eqs. 11a-c, Ch. VIII, see notes at end of chapter.

which is at 2000 persons living in one regional center for most of the world (it is, however, 2500 in the United States).

As an hypothesis for trial, the following definitions for standardized purposes is proposed:

Rank $p^1$	Regional plurel	Representative population = $10^p$		Limits of each plurel	
		( $10^p$ )	(1)	Lower limit = $.2 \times 10^p$	Upper limit = $2 \times 10^p - 1$
	RURAL				
1	Household or centidem	$10^1$	10	2	19
2	Hamlet or decidem	$10^2$	100	20	199
3	Village or dem	$10^3$	1000	200	1999
	URBAN				
4	Town or dekadem	$10^4$	10,000	2000	19,999
5	City or hectodem	$10^5$	100,000	20,000	199,999
6	Metropolis or kilodem	$10^6$	1,000,000	200,000	1,999,999
	NATIONAL				
7	Small nation or myriadem	$10^7$	10,000,000	2,000,000	19,999,999
8	Large nation or macrodem	$10^8$	100,000,000	20,000,000	199,999,999
9	Continental bloc or megadem	$10^9$	1,000,000,000	200,000,000	1,999,999,999
10	WORLD	( $10^{10}$ )		2,000,000,000	(Estimated about 2.2 billion)

The hermit is the limiting case of this series and is outside the definition of a plurel. The "representative population" of each regional plurel is a central tendency in round numbers which need not connote any particular type of mean, or median of its class-interval. It is five times the lower limit of each class and one half the upper limit. The formula for this classification of regional plurels on the basis of multiples of ten (i.e. equal logarithmic increments in size) is:

$$(10^1 P)^p :: L^2, \text{ regional plurels, decimally classified by size} \\ (\text{Eq. 11d, Ch. 8})$$

This formula specifies a series of class-intervals ( $r$ ) of popula-

tion, fixed by a variable exponent on the base of ten persons, and cross-classified with indefinite areas. Each interval of plurel size determines what plurels fall into that plurel-class-interval but also each geographically bounded area determines the population to be counted, so that the dependence between the population size and its single corresponding area is a mutual one as indicated by the double colon.

In discussing this classification of regional plurels, it should be noted that current names and legal definitions of "towns," "city," etc. lack standardization and hence will often overlap into the classes above or below its class in the tabulation above. Also there will be extreme deviations, such as metropolitan New York whose population exceeds two million, which overlap into the distribution of regional plurels which constitutes an adjacent class interval. But with all its exceptions at present, it is submitted that such a classification can standardize terms for regional plurels so that other social scientists can know what order of magnitude is meant when the term "town," "urban plurel," "small nation," etc. is used. As with the metric system, the advantages of adopting a decimal hierarchy of units with precise boundaries may gradually come to outweigh the haphazard historically determined meanings of the terms now used to denote regional plurels of increasing sizes. The formula, Eq. 11d, Ch. 8, discovers no new principle in nature; it imposes an artificial principle upon nature in order to reduce nature to orderly arrangement. It is not an induction towards stating a scientific law; it is rather a proposal to agree upon an operational definition of a set of concepts which provide man with a regular and orderly way of responding to phenomena of this kind (i.e. to names of regional plurels). This is a case of Lundberg's postulate, repeatedly emphasized in his companion volume *Foundations of Sociology*, that in science uniformities (i.e., laws) are fixed not so much by the intrinsic behavior of phenomena as by man's choice of that way of responding to phenomena (such as by an algebraic formula) which is most objective, exact, parsimonious, and fruitful for prediction and control. The essence of this classificatory proposal, Eq. 11d, is the ordinal ranks (defining the population by exponents on the base 10) and not the current terms "hamlet," "village," "town," etc. which are merely used as approximate labels to connect the

new algebraic definition with familiar connotations to the social scientist. If these labels yield confusion between new exact meanings and current inexact meanings any coined term can be substituted such as "demo," defined as one thousand persons living in a contiguous area, with the usual metric prefixes attached. Thus, millidem (= a hermit) =  $10^0$ , centidem (= a household) =  $10^1$ , decidem (= a hamlet) =  $10^2$ , dem (= a village) =  $10^3$ , dekadem (= a town) =  $10^4$ , hectodem (= a city) =  $10^5$ , kilodem (= a metropolis) =  $10^6$ , myriadem (= a small nation) =  $10^7$ , macrodem (= a large nation) =  $10^8$ , megadem (= a continental bloc) =  $10^9$ . Thus a demic characteristic would mean a characteristic pertaining to a regional plurel of between 200 and 2,000 persons, i.e. a term for "village" which has precise limits; a kilodem characteristic pertains to a metropolitan or other regional plurel of between 200,000 and 2,000,000 persons. In S-notation a birth rate expressed as fifteen per thousand persons would be  $\text{dP}$  using the capital letter D as a class-interval script to state the unit, one dem. Similarly a diphtheria death rate of P persons per hundred thousand in some region would be  $\text{HD P}$ , in hectodem units. Whether this proposal to apply the metric system to human populations meets a real need for terms that are more convenient and precise than our present ones is unknown until social scientists have had opportunity to react by using or neglecting these terms.

### C. Indicatory Densities, $IL^{-1}$ , Culture Areas

An indicatory density is the ratio of any characteristic, qualitative or quantitative, to a space (line, area, or volume). The concepts of culture center, culture gradient, culture area, culture margin, regions, or districts, frontiers and pioneer belts, zones of degeneration, regional communities of any kind, all denote data which would be classified by S-formula under indicatory densities, the  $IL^{-1}$  array of the quantic solid.

For example, for the density of the qualitative characteristics of culture areas, see S. 1, Ch. VIII; of languages of commerce, see S. 2, Ch. VIII; of city zones in Jerusalem, see S. 6, Ch. VIII.<sup>9\*</sup> For an example of a density of an all-or-none indicant of free speech, see S. 17, Ch. VIII; for indicants of soot per block, see S. 18, Ch. VIII; for autos per State, see S. 19, Ch. VIII; for cap-

\* For Eqs. 12a-b, see notes at end of chapter.

ital invested by continents, see S. 20, Ch. VIII. For correlated indicatory densities of civilization and latitude, see S. 21, Ch. VIII; and of national independence with language, religion, race, rainfall, health, and illiteracy, see S. 22, Ch. VIII. (Cf. S. 6, Ch. VI.)

*D. Temporal and Lineate Densities,  $TL^{-1}$ ,  $LL^{-1}$*

To complete the survey of societal densities the less frequent densities of time and linear space indices may be noted. Temporal densities occur chiefly: (a) as the frequency within a period of some event or action reduced to spatial units, as in the number of accidents per mile of railroad ( ${}^2T^0L^{-1}$ ), or (b) as a duration per unit of space, as in the inverted density of skull capacity per age of child,  $L^3T^{-1}$  (S. 12, Ch. IX), or the hours per journey from Pittsburgh to Philadelphia,  $TL^{-1}$  (S. 84, Ch. X).

Lineate densities are the number of units of length per unit of space. An example is the kilometrage of railways per square kilometer of territory in the various nations, S. 23, Ch. VIII.<sup>10</sup> \*

Another example is a "topographical index," which might be constructed in order to correlate the ruggedness or mountainousness of a region with some other characteristic such as "love of freedom" measured by the percentage of years of independence in that country's history. For the topographical index a grid of rectangular co-ordinates at standardized intervals could be superposed on a topographical map of specified scale. The number of intersections, made by the grid co-ordinates and the altitude contour lines of the map, per unit of map area would be the topographic index. This index is zero for a level plain and rises as the land surfaces become steeper in hillsides approaching cliffs as the upper limit. Such a topographical index could measure "gradients" of any societal isopleths and might prove useful in many ecological studies.

IV. *S-SITUATIONS*

The notation of S-theory and the specific formulae in this chapter, as in the other chapters, are largely methodological hypotheses. They claim that sociological data can be handled with increased precision, brevity, objectivity, and flexibility by means

\* For Eqs. 13a-e, Ch. VIII, see notes at end of chapter.

of these S-symbols. To test hypotheses evidence is required. A small sample of evidence follows in the twenty-five S-situations which represent the way these S-formulae fit diverse sets of data and reduce them all to orderly classification.

S. 1



Ref.: Cooley, Angell, and Carr, *Introductory Sociology*, Scribners, 1933, p. 84. From Alfred Louis Kroeber's *Anthropology*, Harcourt, Brace and Company.

Descriptive formula:  $S_1 = \underline{L}_1^2$

Quantic number = 0;0;2;0

Legend:

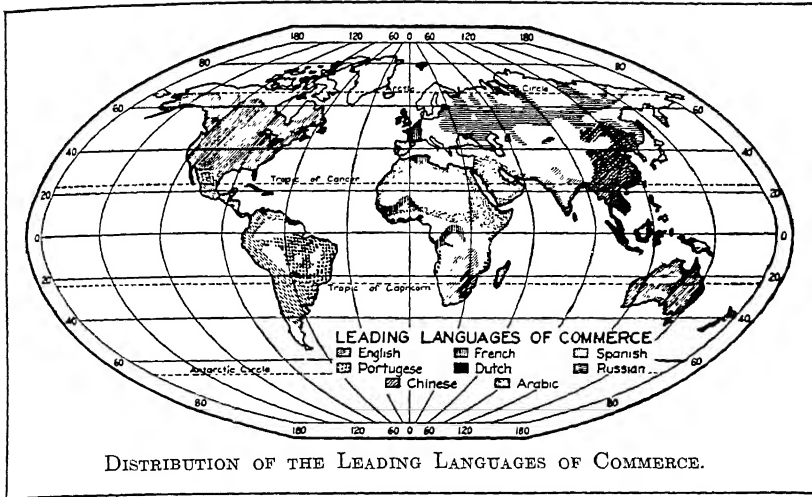
$S_1$  = The situation  
records

$\underline{L}_1^2$  = 15 zones of Amerind culture.

Comment:

A culture area is an area defined by one plurel with a system of characteristics. It combines one physical area and many intercorrelated indicators of culture traits. These fifteen systems of culture traits are represented by the attributes  $L_1^1$ , in the implicit product above,  $I \underline{L}_1^2 = \underline{L}_1^2$  (see Eq. 5, Ch. VIII).

S. 2



Ref.: Huntington and Carlson, *Geographic Basis of Society*, Prentice-Hall, p. 39.

Descriptive formula:  $S_2 = \underline{L}_1^2$

Quantic number = 0;0;2;0

Legend:

$S_2$  = The situation  
records

$\underline{L}_1^2$  = 8 language zones of the world.

S. 3



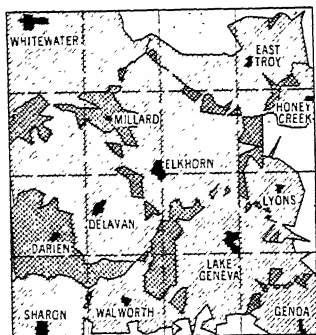
Ref.: Régie des Travaux du Cadastre et d'Amélioration Focier, *Notice sur le Régime Focier et le Cadastre des Etats de Syrie et du Liban*, Beyreuth, Syrie.

Descriptive formula:  $S_3 = L_1^2 : m$ 

Legend:

 $S_3$  = The situation  
records $L_1^2$  = an area  
divided into
 $|_1 = 2$  qualitative portayals  $\left\{ \begin{array}{l} \text{areas} \\ \text{photographed} \\ \text{areas drawn} \\ \text{to scale} \end{array} \right.$   
and into
 $|_m =$  plots

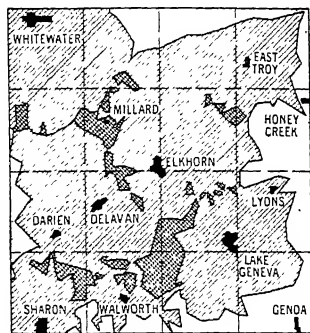
S. 4



- Village or city center.
- Trade at two or more centers.
- Trade at one center.
- Trade outside the county.

**TRADE COMMUNITIES**

Twelve villages and small cities situated in the county serve as trade centres for the farm homes precisely as for the village and city homes, and all the homes trading at the same centre form a trade community. Township lines six miles apart indicate the distance.



- Village or city center.
- Take local papers printed at two or more centers.
- Take local papers printed at one center.
- Take local papers printed outside the county.

**LOCAL-PAPER COMMUNITIES**

The homes taking the same local paper form a paper community. Such communities are pretty well defined and conform closely to the trade and banking communities.

Ref.: Cooley, Angel, and Carr, *Introductory Sociology*, Scribners, 1933, p. 255.

Descriptive formula:  $S_4 = L_1^2 : m$ 

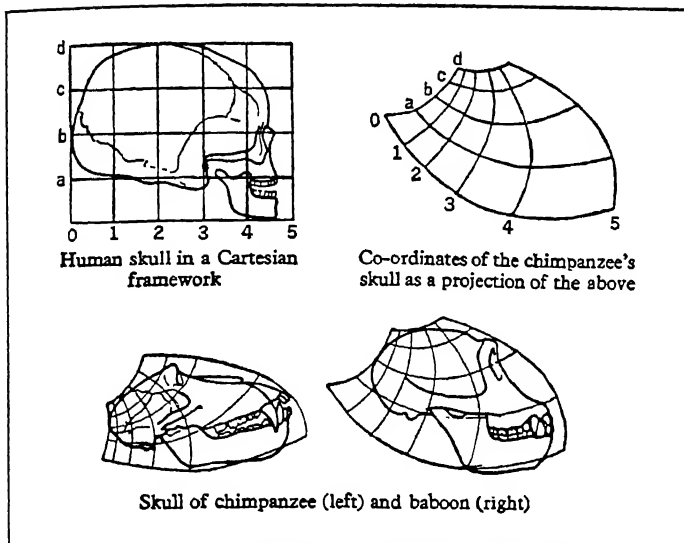
Legend:

 $S_4$  = The situation  
records $L_1^2$  = mapped areas

Quantic number = 0;0;2;0

 $|_1 =$  of 2 kinds  $\left\{ \begin{array}{l} \text{trade and} \\ \text{local paper} \end{array} \right.$   
with
 $|_m = 4$  subdivisions

## S. 5



Ref.: Hogben, Lancelot, *Mathematics for the Millions*, W. W. Norton, p. 450.

Descriptive formula:  $S_5 = L_1^2 : L_1$

Quantic number = 0;0;21;0

Legend:

$S_5$  = The situation  
records

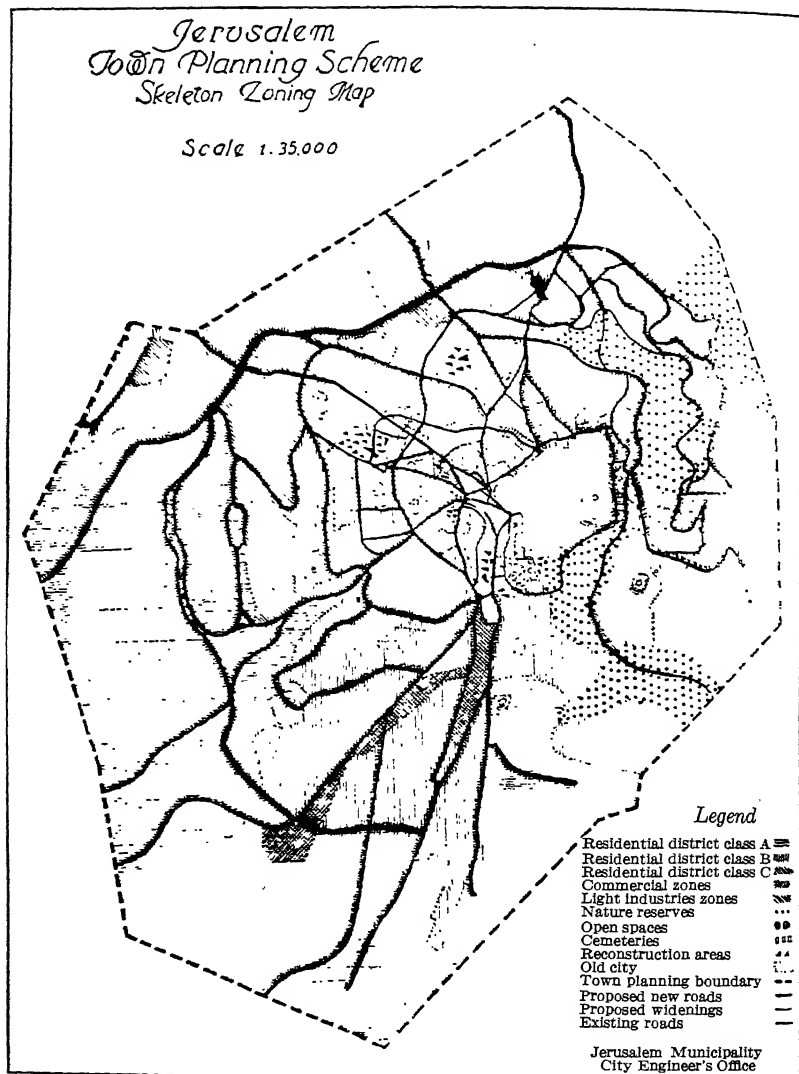
$L^2$  = the projected area

$|_1$  = of 4 kinds of anthropoid skulls  
:= and their corresponding  
 $L_1$  = pattern of co-ordinate lines

Comment on notation:

In analyzing a situation such as this some subjective interpretation seems unavoidable. Should the single indefinite person be written explicitly as  $\angle P$ , or would  $L_1^2$  specified as "area of anthropoid skulls" cover it? The rule is, that if parties are attached to spatial areas that are merely named, their identifying characteristic (the name) is included in the class script on the  $L$ . (See Rule #44.) Should the four skulls be symbolized as four kinds of areas,  $L_1^2$ , or as a series at 4 dates in time,  $T^{-1} : L^2$ ? The evolutionary implication is strong, but the graph as presented merely states 4 types and the analyst must avoid reading into a graph more than is explicitly presented. The resulting quantic classifies the situation as a lineate areal density, a pattern of lines in areas. The density, which strictly is a ratio, is undeveloped, the ingredients for it only being stated. This is reflected in the descriptive formula by positive exponents on both the areal and the linear components, and by a colon, denoting aggregation instead of the juxtaposition denoting multiplication between them.

## S. 6



Ref.: *Social Relationships in the Near East*, American Press, Beirut, Syria, p. 229.

Descriptive formula:  $S_6 = \underline{L}_1^2 : \underline{L}_1 : \underline{m}$

Quantic number = 0;0;21;0

Legend:

$S_6$  = The situation

$|\underline{m}$  = subsections of each

is a record of

and corresponding

$L^2$  = the area of Jerusalem town-  
ship

$L_1^{+1} = 4$  classes of lines (streets,  
boundaries, etc.)

subdivided into

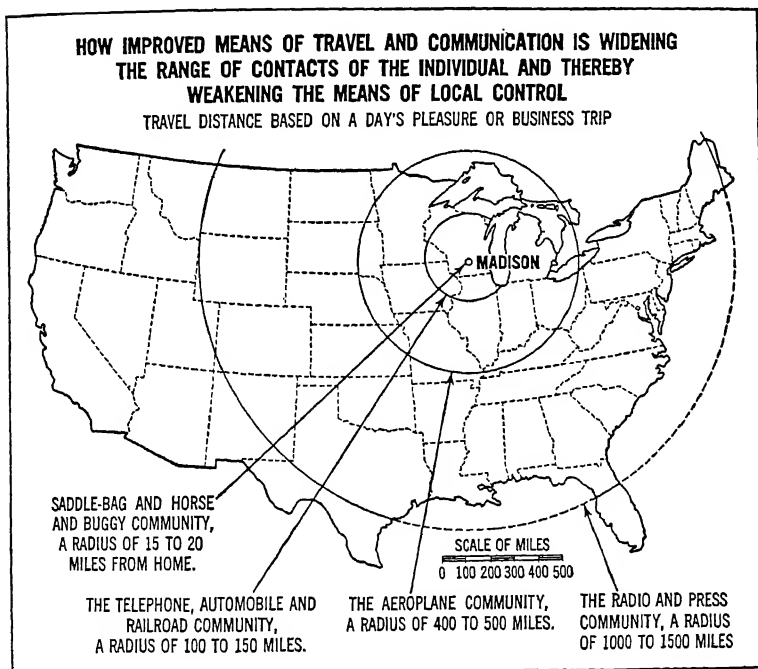
with

$|\underline{l}_1 = 10$  zones

$|\underline{m}$  = a number of examples of each

with

### S. 7



Ref.: Gillin, Dittner, and Colbert, *Social Problems*, The Century Co., 1928, p. 35.

Descriptive formula:  $S_7 = \underline{P} : \underline{L}_1^2 : \underline{m}$

Quantic number = 0;0;2;1

Legend:

$S_7$  = The situation

$\underline{L}_1^2 = 4$  concentric zones of commu-  
nication,

records for

and

$\underline{P}$  = Madison, Wis.

$|\underline{m}$  = the 48 State areas are also  
mapped



Descriptive formula:  $S_8 = \angle P : \underline{L}_1^2 : m : n$

Quantic number = 0;0;2;1

Legend:

$S_8$  = The situation

$|_1$  = subareas

records for

and

$\angle P$  = one unspecified person (i.e., a type)

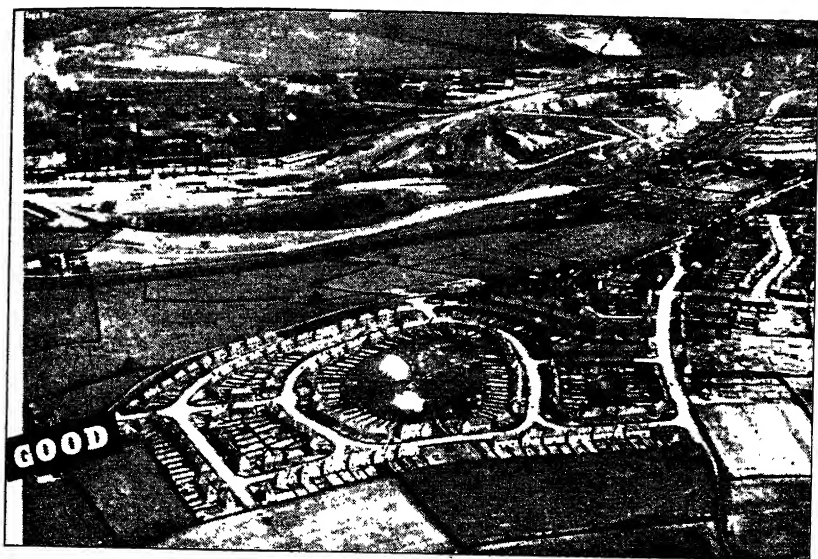
$|_m$  = subcenters

$|_.$  = compared with

$\underline{L}^2$  = an area of brain surface subdivided into

$|_n$  = phrenological areas

## S. 9



Ref.: *Life*, May 23, 1938, p. 59.

Descriptive formula:  $S_9 = I^0 : \underline{L}_1^2 : m : n : \underline{P}$

Quantic number = 0;0;2;1

Legend:

$S_9$  = The situation,

$|_1$  = work and residence areas subdivided into

$I^0$  = recorded with the attribute "Good," (with respect to housing conditions)

$|_m$  = an indefinite number of blocks subdivided into

has

$\underline{L}^2$  = a photographed area of England

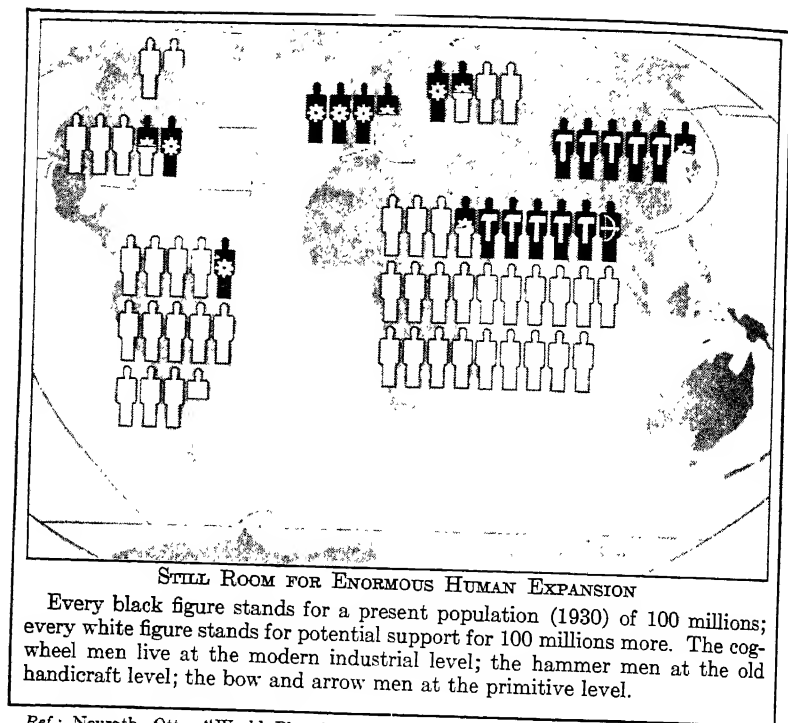
$|_n$  = an indefinite number of house-plots

subdivided into

each with

$\underline{P}$  = a family plurel

## S. 10



Ref.: Neurath, Otto, "World Planning and the U.S.A.," *The Survey*, Vol. LXVII, No. 11, Mar. 1, 1932.

Descriptive formula:  $S_{10} = 'T^0 : \underline{L}_1^2 : P_p : q$

Legend:

$S_{10}$  = The situation

records

$'T^0$  = for 1930

for each of

$\underline{L}_1 = 7$

$\underline{L}^2$  = mapped regions (areas not stated)

Quantic number =  $0;0;2;1$

$:$  = corresponding

$P$  = population

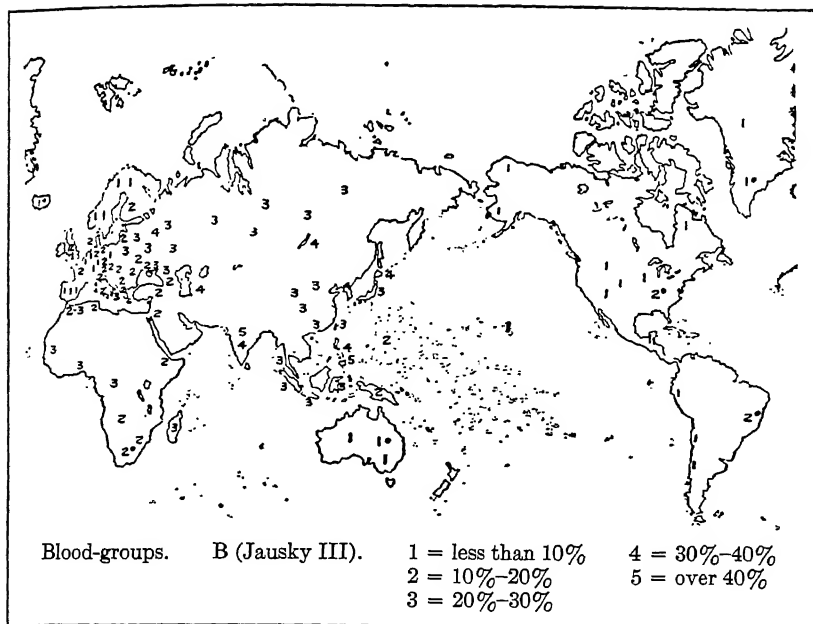
divided into

$|_p$  = possible and actual

subdivided into

$|_q = 3$  occupational level plurels

## S. 11



Ref.: Davis, Allison, "Distribution of Blood Groups and Its Bearing on the Concept of Race," *Sociological Review*, Vol. XXVII, No. 1, Jan., 1935, p. 30.

Descriptive formula:  $S_{11} = {}^1\bar{L}^2 : \%P_p$

Quantic number = 0;0;2;1

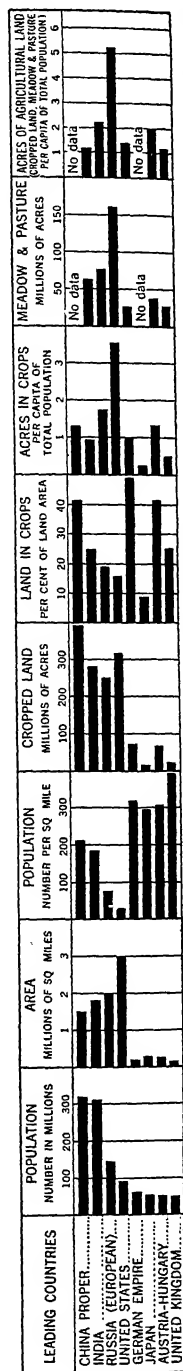
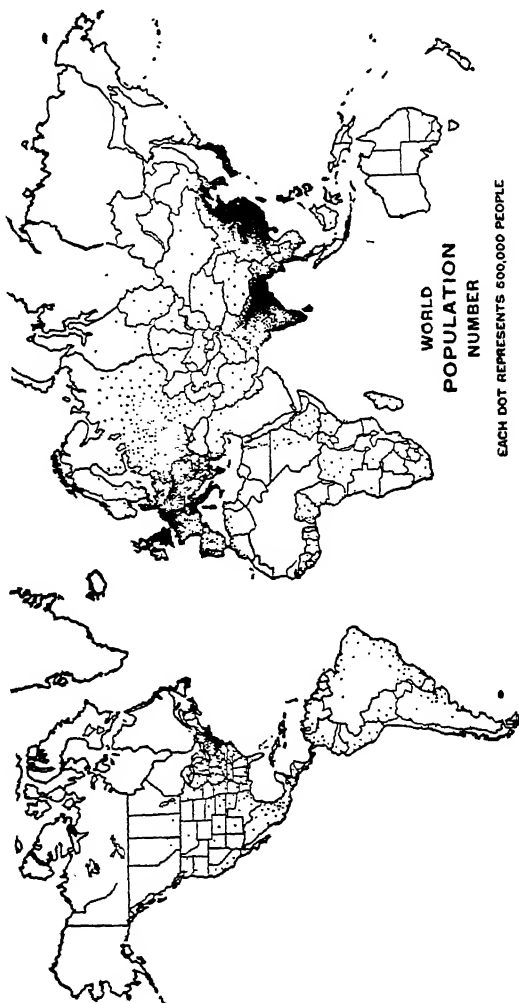
Legend:

$S_{11}$  = The situation  
records

$\%P$  = blood group percentages  
in

${}^1\bar{L}^2$  = points on the earth  
:= with corresponding

$|_p$  = 2 plurels—European and non-European



*Ref.: Reinhardt and Davies, Principles and Methods of Sociology, Proutice-Hall, N. Y., 1932, pp. 374-375.*

## S. 12 (see page 474)

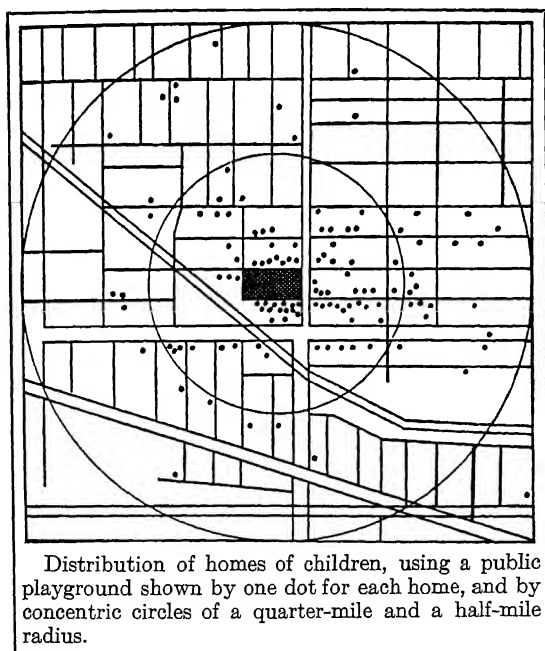
Descriptive formula:  $S_{12} = 'T^0 : L^{-2}P_1 : m$ 

Quantic number = 0;0;2;1

Legend:

 $S_{12}$  = The situation  
records $PL^{-2}$  = the density  
of $T^0$  = at a specific date $|_1$  = the countries of the world  
for each of $P$  = the population, $L^2$  = the area,  
and $|_m$  = 4 classes of land by crops

## S.13

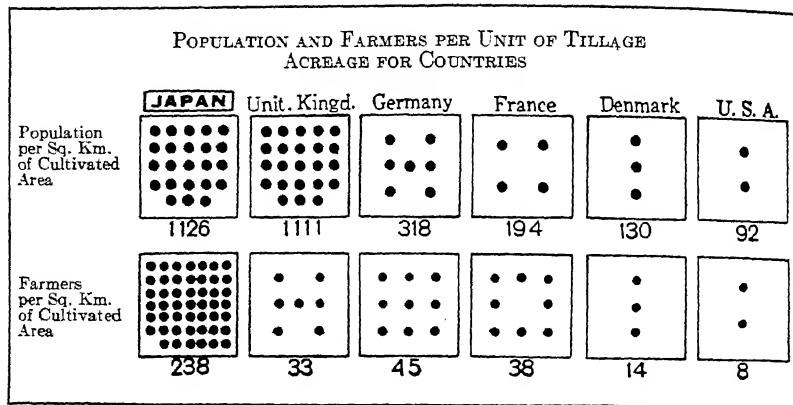
Ref.: White, R. Clyde, *Social Statistics*, Harper and Brothers, N. Y., 1933, p. 184.Descriptive formula:  $S_{13} = ({}_pP : 'L^2)_1$ 

Quantic number = 0;0;2;1

Legend:

 $S_{13}$  = The situation  
records $'L^2$  = home locations  
mapped in ${}_pP$  = the families of playground  
children  
and their corresponding $|_1$  = 2 regional  
classes { blocks  
and  
zones

## S. 14



Ref.: Yano, T. and Shinasaki, K., *Nippon, A Chartered Survey of Japan*, Kikusei-Sha. The First National Building, Kyoboshi, Tokyo, Chart 32, p. 85.

Descriptive formula:  $S_{14} = \frac{P}{L} : q$

Legend:

$S_{14}$  = The situation

records

$\frac{P}{L}$  = the unstated number of persons

$L^{-2}$  = per square kilometer

Quantic number = 0;0;8;1

in each of

$|_p$  = 6 nations

and for

$|_q$  = farming and total populations

## S. 15

*SHOWING NUMBER OF ROOMS WITH SPECIFIED NUMBER  
OF OCCUPANTS, OR PARTS OF OCCUPANTS PER ROOM*

NUMBER OF OCCUPANTS	NUMBER OF ROOMS				
	ONE ROOM	TWO ROOMS	THREE ROOMS	FOUR ROOMS	
1 AND LESS THAN 2					
2 AND LESS THAN 3					
3 AND LESS THAN 4					
4 OR OVER					

② Oh, beware what statistics have done to my anatomy  
And then get some part to make a tubercular case

ENTER

Ref.: Stein, Clarence S., *The Housing Crisis in New York*, Survey, Vol. XLIV, No. 19, Sept. 1920, p. 662.

Descriptive formula:  $S_{15} = {}_p(PL^{-2}) : 1$

Legend:

$S_{15}$  = The situation

records

$p$  = 4 sizes of plurals

$P$  = the number of persons

Quantic number = 0;0;8;1

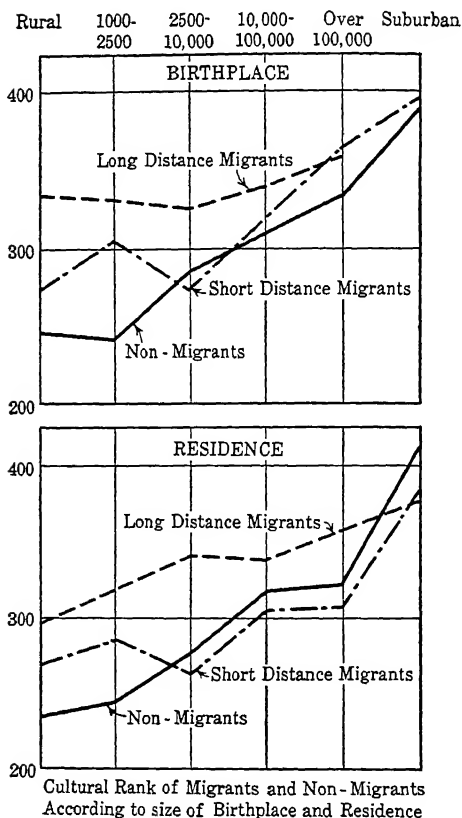
$L^{-2}$  = per room

| : 1 = with the corresponding number of rooms

Comment:

A room might be interpreted as a volume ( $L^3$ ) instead of an area ( $L^2$ ) as here. The area of floor space, however, is the more usual unit in studying crowding and hence a "room" is so symbolized here.

## S. 16



Ref.: Huntington, Ellsworth, *The Selective Action of Migration*, New Haven, Conn., p. 509.

Descriptive formula:  $S_{16} = {}_1L : {}_pP : I$

Quantic number = 0;1;1;1

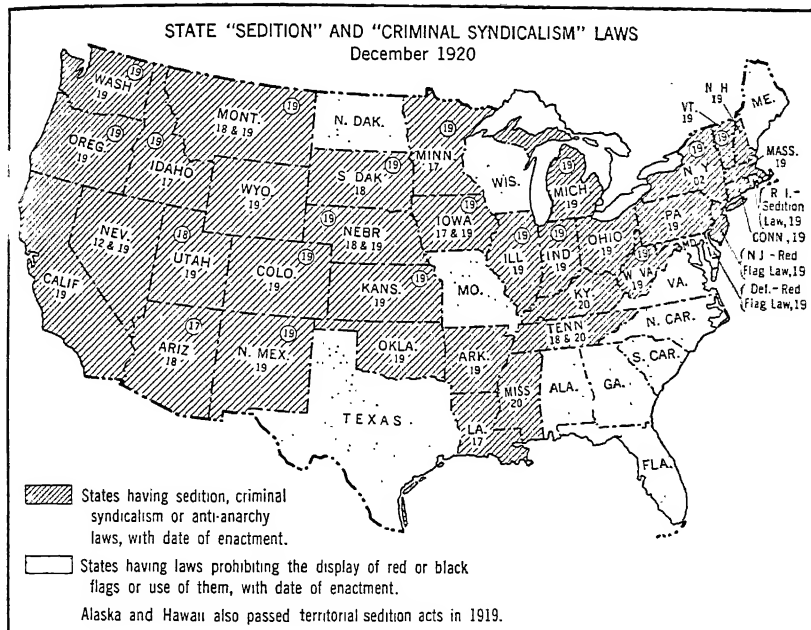
Legend:

$S_{16}$  = The situation  
records  
for each of

${}_1L$  = 3 distances of migration  $\left\{ \begin{array}{l} \text{long} \\ \text{short} \\ \text{zero} \end{array} \right.$

$P$  = a plurel  
from  
 ${}_p|$  = 6 sizes of cities  
each with  
 $I$  = a cultural rank

## S. 17



## THE WAVE OF "SEDITION" LAWS

This map, just issued by the American Civil Liberties Union, shows the extent to which legislation restricting free speech has spread in the United States. These laws with few exceptions have been passed within the last five years. In announcing its campaign for the elimination of these laws the union calls attention to their use to punish persons for expressing unorthodox economic or political beliefs.

Ref.: *Survey*, Vol. XLV, No. 18, p. 623.

Descriptive formula:  $S_{17} = L_1^2:1,0I$

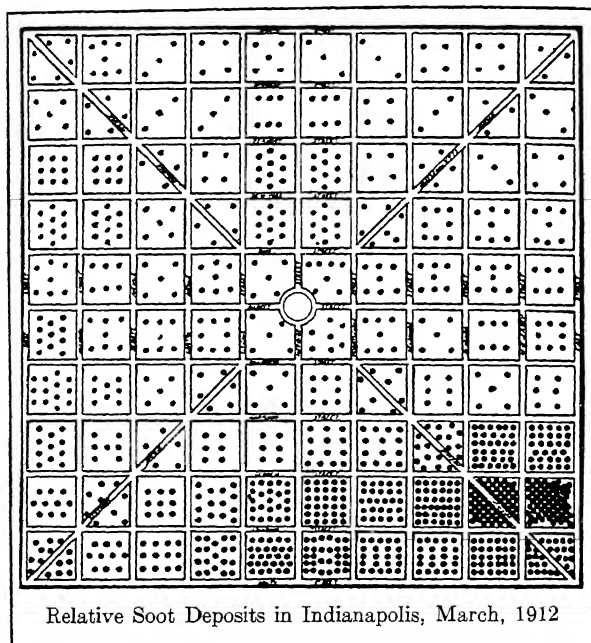
Legend:

$S_{17}$  = The situation  
records  
for each of  
 $|_1$  = the 48 States

Quantic number =  $0;1;2;0$

$L^2$  = whose areas are mapped  
 $1,0I$  = an all-or-none indicant of free speech

## S. 18



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 245. William D. McAbber in the *Survey*.

Descriptive formula:  $S_{18} = {}^T T^0 : \{L\}^2 : I$

Quantic number = 0;1;2;0

Legend:

$S_{18}$  = The situation

in

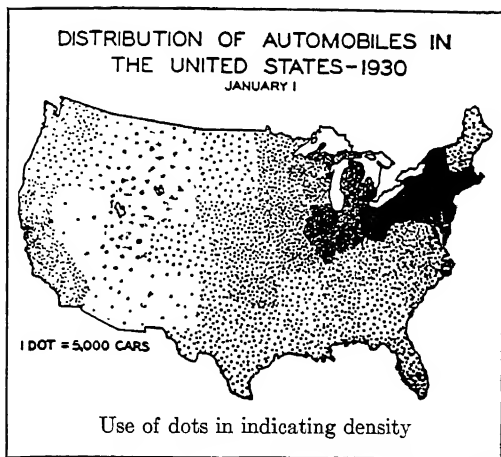
records

$\{L\}^2$  = the blocks of Indianapolis

${}^T T^0$  = for March, 1912

I = the soot deposit

## S. 19



Ref.: Riggeman, R. John, *Graphic Methods for Business Statistics*, McGraw-Hill Book Co., N. Y., 1936, p. 184.

*Descriptive formula:*  $S_{19} = 'T^0 : \underline{L}_1^2 : I$

*Legend:*

$S_{19}$  = The situation

records

' $T^0$  = on a specified date

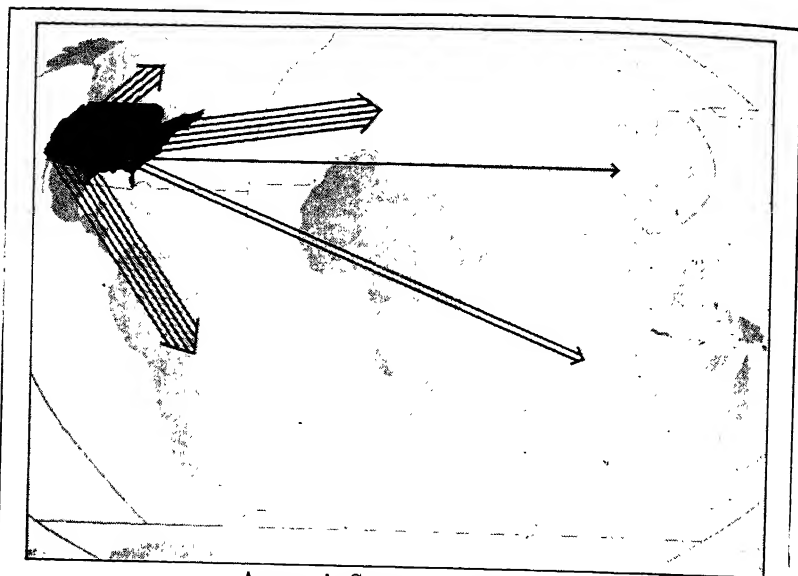
*Quantic number* = 0;1;2;0

for each of the

$\underline{L}_1^2$  = mapped U.S. States

I = the number of autos

## S. 20



AMERICA'S STREAM OF CAPITAL

Every arrow line stands for a billion dollars in long-term credits (including government bonds) representing United States investments in foreign countries in 1931.

Ref.: Neurath, Otto, "World Planning in the U. S. A.," *Survey*, Vol. LXVII, No. 11, March 1, 1932, p. 624.

Descriptive formula:  $S_{20} = 'T^0 : \underline{L}_1^2 : I$

Legend:

$S_{20}$  = The situation  
records

$'T^0$  = for 1931

for each of

Quantic number = 0;1;2;0

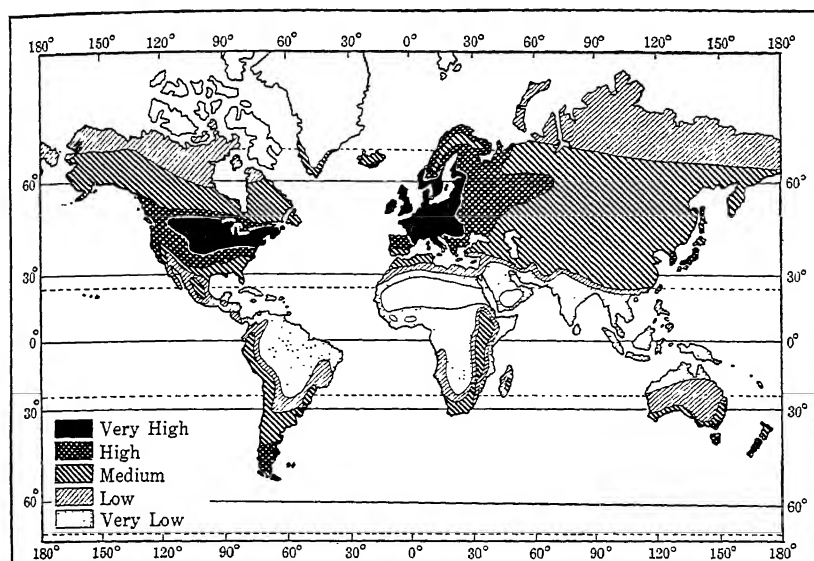
$\{L\}_1 = 5$

$\underline{L}_1^2$  = mapped areas

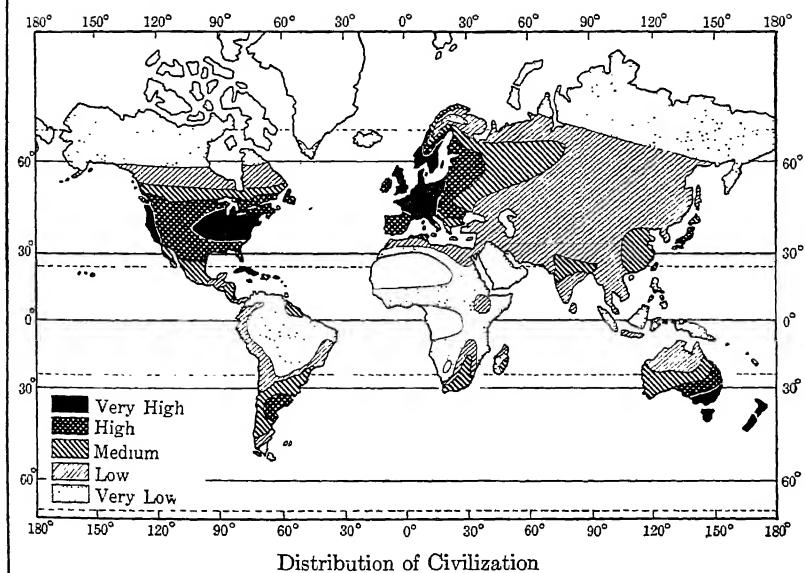
: = a corresponding

$I$  = investment of American dollars

## S. 21



The Distribution of Human Health and Energy on the Basis of Climate



*Descriptive formula:*  $S_{21} = ({}_iI :: {}^iI : \underline{L}^2)_i$

*Quantic number* = 0;2;2;0

*Legend:*

$S_{21}$  = The situation

on each of

records

for each of

${}_iI$  = the temperature zones

cross-classified with

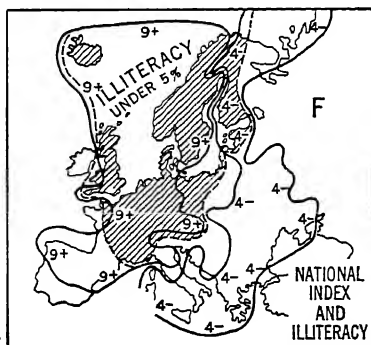
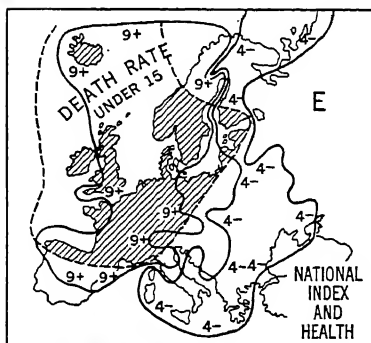
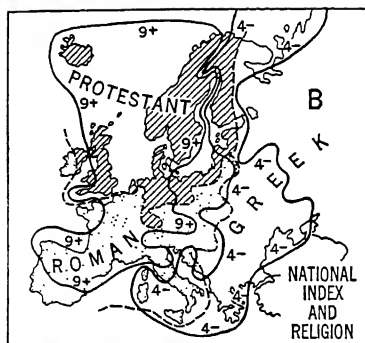
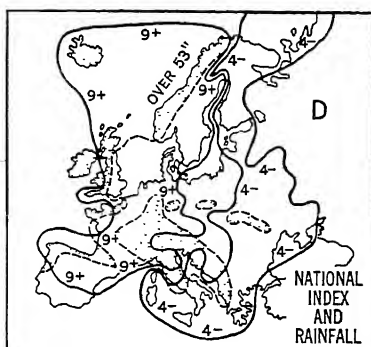
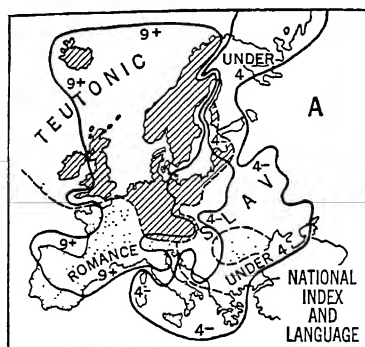
${}^i|$  = 5 rank points

$I_i = 2$  indicants  $\left\{ \begin{array}{l} \text{energy,} \\ \text{civili-} \\ \text{zation} \end{array} \right.$

the corresponding

$\underline{L}^2$  = mapped areas of the world

## S. 22



Correlations of national-index in Europe with language, religion, race, rainfall, health, and illiteracy. In all cases the area inclosed by the isopleth 9+ has a high national-index (i.e., lengthy self-government) and by isopleth 4- has a low national-index. (E, partly after Huntington; F, partly after Hettner.)

Descriptive formula:  $S_{22} = ({}_1I :: {}_i(I) : L^2)_i$

Quantic number = 0;2;2;1

Legend:

$S_{22}$  = The situation

records

${}_1|$  = for every class-interval of

$I$  = an indicant of nationhood

$::$  = cross-classified

with each of

${}_i|$  = the class-intervals

of each of

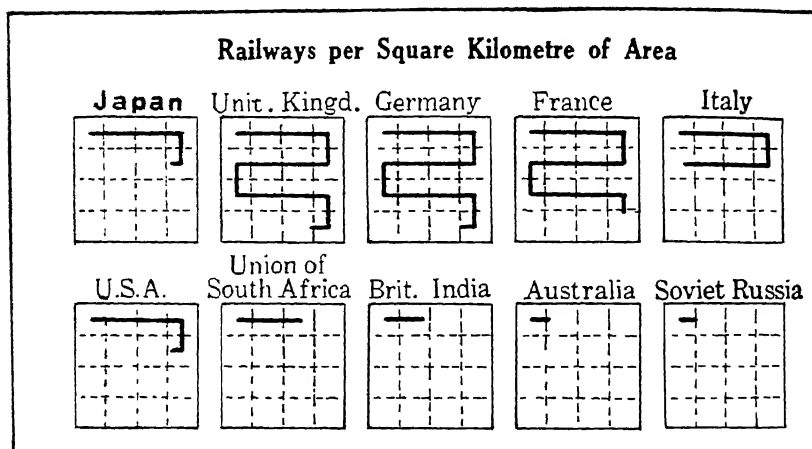
$(I)_i$  = 6 indices of environment

with corresponding

$L^2$  = mapped zones of Europe.

$$({}_1I)_i = \begin{cases} I^0 = \text{language} \\ I_{,,} = \text{religion} \\ I_{,,,} = \text{race} \\ I_{,v} = \text{rainfall} \\ {}_cP = \text{death rate} \\ {}_vP_{,,} = \text{illiteracy} \end{cases}$$

### S. 23



Ref.: Yano, T. and Shinasaki, *Nippon, A Chartered Survey of Japan*, Kukusei-Sha, Tokyo, Chart 135, p. 345.

Descriptive formula:  $S_{23} = (\underline{L}\underline{L}^{-2})_1$

Quantic number = 0;0;18;0

Legend:

$S_{23}$  = The situation

records

$\underline{L}$  = the railway trackage

$\underline{L}^{-2}$  = per sq. kilometer of area

for each of

${}_1|$  = 10 nations

Comment on notation:

The algebraic sum of the exponents  $(-2 + 1)$  is  $-1$ , but the quantic digits, 18, preserve the structure in the situation, in asserting it to be a ratio of lines to areas.

## S. 24



Lighted honeycombs of New York City's industry, finance, and gay living make pretty pictures by night. But actual living facilities of this wealthy city compare unfavorably by daylight with many little rural towns.

Ref.: *Life*, May 23, 1938, p. 57. (Reproduced by permission of Ewing Galloway, N. Y.).

Descriptive formula:  $S_{24} = I^0 : \underline{1} : \underline{m} \underline{L}^3$

Quantic number = 0;0;3;0

Legend:

$S_{24}$  = The situation

and

is a record

corresponding to

$\underline{m}$  = rooms

$I^0$  = the attribute "Bad"

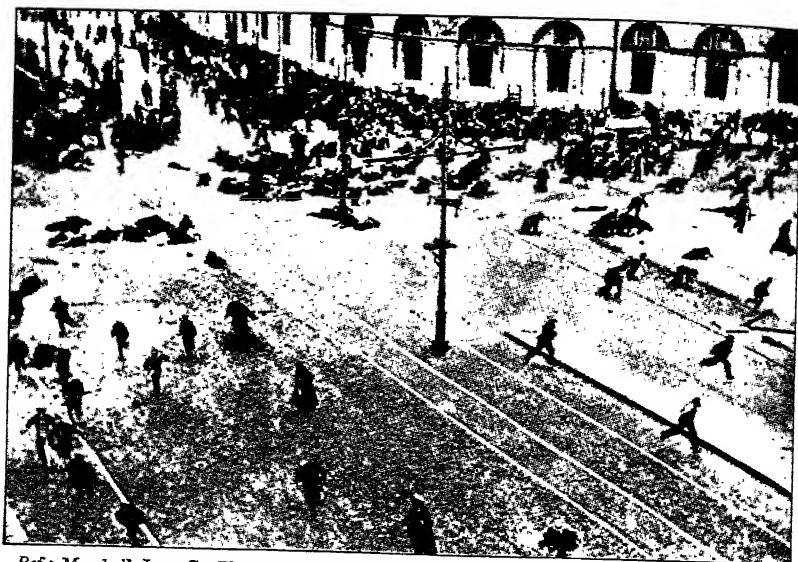
in

of

$\underline{L}^3$  = a photo of New York City (in three dimensions)

$\underline{1}$  = an indefinite number of buildings

## S. 25



Ref.: Marshall, Leon C., *The Story of Human Progress*, Macmillan, N. Y., 1925, p. 502.

Descriptive formula:  $S_{25} = 'T^0 : 'L^3 : ^P P$

Legend:

Quantic number = 0;0;3;1

$S_{25}$  = The photographed situation  
records

$'L^3$  = a particular point in space (the  
machine gun station)

$'T^0$  = a particular moment  
and

: = with a corresponding

$^P P$  = crowd of people

Comment:

In this picture a street crowd is being dispersed by machine gun fire from the roof of the building shown.

The picture imperfectly records a crowd in process of becoming a "unified" crowd as their behavior becomes determined by a single stimulus at one point in space. As they realize the location and effect of the machine gun station, they run for safety from it.

## IV. NOTES

1. For a fuller discussion of the fundamental postulates involved in dealing with geographic space as part of societal situations, see Chapter XII on Human Ecology in Lundberg's companion volume *Foundations of Sociology*.

2. The detailed formulae for these exponents and their attendant descriptives are:

$'L^0$  = a named point

(Eq. 1a, Ch. VIII)

$'L^0$  = an aggregation of named points

(Eq. 1b, Ch. VIII)

- $L_1^{+1}$  = an aggregation or map of lines of 1 qualitatively different classes (Eq. 2a, Ch. VIII)
- $l_1^{+1}$  = a line subdivided into 1 equal sects or units (Eq. 2b, Ch. VIII)
- $l_1^{+1}$  = 1 points on a line (Eq. 2c, Ch. VIII)
- $L_1^2$  = an aggregation of areas, a map subdivided into different regions. (See S. 1 and 2, Ch. VIII.) (Eq. 3a, Ch. VIII)
- $L_{1:m:n}^2$  = a regional hierarchy, as of continents ( $|_1$ ) subdivided into countries ( $|_m$ ), further subdivided into provinces ( $|_n$ ). (See S. 3 and 12, Ch. VIII for  $L^{-2}$ .) (Eq. 3b, Ch. VIII)
- $l_1^2$  = successive equal class-intervals, 1 in number, of areas as in a distribution of houses by number of rooms. (See S. 15, Ch. VIII.) (Eq. 3c, Ch. VIII)
- $l_1^2$  = 1 points in some area (Eq. 3d, Ch. VIII)
- $L_1^3$  = an aggregation of three-dimensional regions, such as air lanes, different rooms of a house, etc. (Eq. 4a, Ch. VIII)
- $l_1^3$  = successive equal class-intervals, 1 in number, of volume, as in a frequency distribution of containers by capacity (Eq. 4b, Ch. VIII)
- $l_1^3$  = 1 points in some three-dimensional region (Eq. 4c, Ch. VIII)

In the great majority of recorded societal situations the exponents as described above serve to analyze the situation. Occasionally, however, the standard deviation of a set of areas, or the number of spaces, regardless of their magnitudes, will be recorded. In such situations the exponents denote the statistical moments, as well as denoting the physical dimensionality of the space as point, line, area, or volume. Here there is a double use of exponents which should be noted to avoid confusion. The first and commonest use of spatial exponents is to state the number of physical dimensions—0 for a point, 1 for a line, 2 for an area, 3 for a volume. The second and rarer use of spatial exponents is to state the statistical moment—1 for a mean, 2 for a variance, 3 for skewness, etc. This second meaning is the usual meaning of the indicatory exponents. When the two uses of the exponent occur in the same situation, there are two notational alternatives as to ways of writing the quantic number and classifying the situation. One way is to consider the  $L^1$  with its exponent denoting physical dimensions as an index ( $I$ ), and then apply the operations called for by any further exponents denoting the statistical moments, as in the indicatory sector. The other and better way is to recognize two exponents and write both their digits in the quantic number. Thus a variance of areas would be:  $\sigma^2 = \Sigma (aL^2)^2 / |_1$  ( $aL^2$  being in units of deviation from the mean area) and its quantic would be 0;0;22;0. The "statistical" exponent may be written first, and the "physical" one second. Again the number of areas would be  $(L^2)^{\sigma_{21}} = |_{21}$ , and its quantic, 0;0;02;0. The first quantic digit in a sector determines the class in the quantic classification. (The double quantic digit here is a little different from its use in a homosectoral ratio where, in strictness, when the exponents are subtracted, their quantic digits should be separated by a minus sign. No confusion is likely, however, in omitting the minus sign as in the quantic number of 0;0;18;0 for  $L^{+1}L^{-2}$ , illustrated in S. 23, Ch. VIII.)

This doubled exponent with the physical and the statistical meanings is systematically classified in its dynamic aspects in Chapter X in discussing spatial processes.

3. When the number of units of length or area is not explicitly stated (as in the ordinary map showing relative areas only), the  $L$  is underlined to denote indefiniteness of amount. If the situation does not even comprise a map but merely names a region or states a plurel or characteristic of some region, the  $L$  is not written as a base letter, but is included as a qualitative characteristic in the descripts of the  $P$  or  $I$ . Cf. S. 12, Ch. VIII with S. 4, Ch. VIII and with S. 6, Ch. V.

4. Of course spatial relations, though usually expressed as ratios which are here our definition of density, may also be expressed as a product in such indices as ton-miles, passenger-miles, etc. (See S. 59, Ch. X.)

5. 1 hectare =  $100 \times 100$  meters = .01 square kilometers = 2.471 acres = .003861 square miles. 1 square mile = 259 hectares = 2.59 square kilometers.

6. This Eq. 8, Ch. VIII has methodological significance in that it suggests a way to convert ordinal indicants to cardinal units. By using density as a criterion, why not boldly *define* the successive equal steps in environmental utilization by means of Eq. 8, as "those productive-cultures which support densities of population as given by the exponent of Eq. 8a." This criterion is objective and precisely determinable. It specifies in operational terms the types of culture to be taken as standard equidistant points on the scale of "environmental utilization." Note that this criterion places Miss Semple's mixed category, "agriculture with some industry," as intermediate between points 4 and 5. This illustrates how this method of scale construction would locate any given productive-culture. (S. 12 and 14, Ch. VIII provide some relevant data.)

This proposal for constructing a cardinal scale of productive-cultures needs exploration to qualify and delimit it more carefully. Thus, for example, it may be expected to be less satisfactory for pioneer belts than for long-settled districts. Furthermore, the size of district observed needs standardizing, or else one square mile of a little village may show a spuriously higher density than the million square miles of an industrialized nation.

7. These are symbolized by the usual rules of S-notation as:

$$(PL^{-2})_p = \text{density of plurels} \quad (\text{Eq. 9a, Ch. VIII})$$

This states an index of density of population for each of  $p$  plurels.

$$_p(PL^{-2}) : _p = \text{plurels of density, i.e., rural and urban plurels} \quad (\text{Eq. 9b, Ch. VIII})$$

This states that for each of  $_p$  successive class-intervals of a density index there is a corresponding plurel,  $| : _p$ . In the dichotomy, "rural-urban,"  $_p$  is 2.

$$_p(PL^{-2}) : _p = \text{a plurel of density, i.e., a rural or an urban plurel} \\ (\text{Eq. 9c, Ch. VIII})$$

This specifies one plurel defined by a particular class-interval of the density index, such as "the rural population."

${}_p(PL^{-2}) : {}_p =$  plurels of one density, i.e., either rural plurels, or urban plurels  
(Eq. 9d, Ch. VIII)

This specifies all the plurels,  ${}_p$  in number, which correspond to (i.e., are defined by) a particular class-interval of the density index.

The notational logic of Eq. 9b, c, and d is the same as in the case of distributions of plurels-by-size. (Eq. 4, Ch. V.)

8. These three classes of crowds may be symbolized as follows:

$${}^pP : I^0 : {}^pL^3, \text{ a scattered crowd} \quad (\text{Eq. 11a, Ch. VIII})$$

This asserts  ${}_p|$  persons, each with a stimulus  $I^0$  in some point  ${}^pL^3$  of space.

$$P_p : I^0 : {}^pL^3, \text{ a nucleated crowd} \quad (\text{Eq. 11b, Ch. VIII})$$

This asserts  ${}_p$  plurels, each with a stimulus at some point of space, i.e., the number of stimuli and points is  ${}_p$ .

$$P, : I^0 : {}^pL^3, \text{ a unified crowd} \quad (\text{Eq. 11c, Ch. VIII})$$

This asserts a single plurel with one corresponding stimulus at one point of space.

In case the stimulus, instead of being an impersonal one  $I^0$ , is a person or plurel,  ${}^pP$ , replaces the  $I^0$  in Eq. 11, and the crowd ceases to be a plurel and becomes a group, for it now involves interrelation of people as shown by its quantic number  $t : i ; 3 ; 2$ . Also, fuller description of crowd phenomena requires the inclusion of the dynamic aspects, reserved for later chapters. (For a classification of crowds similar to this one, see Ref. 49.)

9. Note that these qualitative densities are attribute-space products or ratios. Thus in S. 2, Ch. VIII, the languages are attributes, so that  $L_1^2 = I_1^0 L_2^2$ . (See Eq. 5, Ch. VIII.) Note further that the arithmetic value of any quantity with a zero exponent is unity, so that  $I^0 = 1$ . Whether a quantity such as  $L^2$  is multiplied by one or divided by one, the result is the same. Hence we may interchangeably speak of  $I_1^0 L^2$  as a qualitative product, an area of each language, or as a qualitative ratio,  $L^2/I_1^0$ , the area per each language. The ratio of one quantity to one quality is equal to the product of the quantity and the quality:

$$(I, I_1^0 = (I,)/I_1^0 = (I,,) \text{ the equality of a qualitative product and a qualitative ratio} \quad (\text{Eq. 12a, Ch. VIII})$$

A qualitative density,  $I^0 L^{-1}$ , is simply a space characterized by that quality. Any regions on a map, such as the various nations conventionally shown in various colors, are such qualitative densities. Each region has one quality and but one degree or amount of it. Thus any classification of space, denoted by  $L_1^1$ , can be re-expressed in the form of qualitative densities:

$$L_1^1 = I_1^0 L^1 = (I^0 L^1)_1 = (L^1/I^0)_1 \quad (\text{Eq. 12b, Ch. VIII})$$

10. The following types of spatial densities are also possible:

1. Punctual linear densities,  $\Sigma {}^1L^0 L^{-1}$ , the frequency of points on a line  
(Eq. 13a, Ch. VIII)
2. Punctual areal and cubic densities,  $\Sigma {}^1L^0 L^{-2}$ ,  $\Sigma {}^1L^0 L^{-3}$ , of points in a plane, or clusters of points in space  
(Eq. 13b, Ch. VIII)

3. Lineate linear densities,  $LL^{-1}$ , as the percentage of paved stretches of a roadway (Eq. 13c, Ch. VIII)
4. Lineate cubic densities,  $LL^{-3}$ , as of wires in some volume (Eq. 13d, Ch. VIII)
5. Areal cubic densities,  $L^2 L^{-3}$ , of planes in some volume, as in shelf-area per cubic foot of an icebox, or the pages of a volume (Eq. 13e, Ch. VIII)

# *PART V*

## THE TIME SECTOR, ${}^tT_t$

*studying situations defined by  $T^{1,2}$ ;  $I^i$ ;  $L^l$ ;  $P^p$*



## Chapter IX

### DURATIONS, $T^{+1}$

#### I. THE DIMENSIONS OF TIME, $T_t^t$

In the preceding Parts II, III, and IV the static aspects of the indicatory, populational, and spatial sectors of societal situations have been successively expounded. Their dynamic aspects are now taken up by considering situations in which time is explicitly involved. Chapters III through VIII inclusive have dealt with phenomena for which the temporal digit of the quantic number is zero—the  $T^0$  array in the quantic solid S. 33, Ch. II. Consider now phenomena classified under the quantic temporal digits other than zero, the  $T^{\pm 1}$  and the  $T^{\pm 2}$  arrays of the quantic solid.

##### A. The Temporal Exponent, $|^t$

Static, timeless, and instantaneous phenomena are symbolized by an exponent of zero on the time index,  $T^0$ .<sup>1\*</sup>

Time ordinarily occurs in societal situations either as a duration, or as a rate of change (an amount of change divided by the time in which it occurred).

$T^{+1}$  = a duration, a period of time (Eq. 2a, Ch. IX)

(I) $T^{-1}$  = a velocity of change, an index per unit of time<sup>2†</sup>  
(Eq. 3a, Ch. IX)

Duration and change are but two aspects, with differing emphasis, of the same phenomenon of something existing in time. Any finite duration ends, whether gradually or suddenly, and thus changes sometime. A long duration can be considered as a slow velocity of change; a short duration as swifter change. Both are primary indices, i.e., indices with an exponent of 1. The distinction between the exponent of plus 1 and minus 1 is useful, however, even though the boundary is sometimes somewhat artificial. All ages of persons, or things, at a given moment, and all aggregations or series of durations from variable initial dates up

\* For Eqs. 1a-c Ch. IX, see notes at end of chapter.

† For 2b-f and 3b-f, Ch. IX, see notes at end of chapter.

to a common present moment, are considered as durations and symbolized by  $T^{+1}$ . All changes, velocities, processes, rates of growth, from a common initial date to either a common or variable terminal dates, are considered as changes and are symbolized by  $T^{-1}$ . A population pyramid is a distribution of people by durations (ages), but if one set of persons were counted year after year as they grew up and died off, their frequency distribution at each age would be a velocity of dying, and the exponent would be minus one. (See S. 3, Ch. IX for this distinction.) A definition of duration and of change, which is more exact and more amplified than the verbal statement above, is given by the fifteen graphed situations at the end of this chapter, all of which involve durations, and the eighty-five graphed situations at the end of Chapter X, all of which involve change. The situations S. 74, 75, 77, 79, and 85, Ch. X illustrate both, as these situations record changing durations, such as population pyramids changing from census date to census date. The present chapter covers durational situations, while the next chapter on Change takes up societal processes, growth, and change generally.

Whenever a velocity of change is speeded up or slowed down there is acceleration—a change of pace. This is measured by dividing the amount of the difference in the velocities of two periods by the time interval in which that difference developed. This ratio of a velocity per unit of time, or the rate of change of velocity, means that the original societal change has been twice divided by a time index. An acceleration is thus defined by the presence of time to the minus two in the quantic formula:

$$T^{-2}I = \text{an acceleration of change of any index} \quad (\text{Eq. 4a, Ch. IX})^{3*}$$

$$|^s = 8j;l;p$$

Acceleration is the essence of a force, whether physical or societal, since a force is defined as that which accelerates change in something. (See S. 21, Ch. II.) This class of temporal phenomena will be treated in Chapter XI, which comprises secondary temporal phenomena defined by  $T^{-2}$ .<sup>4</sup> †

As usual in S-theory, the exponent symbolizes the operational

\* For Eqs. 4b-6, Ch. IX, see notes at end of chapter.

† For Eq. 7, Ch. IX, see notes at end of chapter.

degree to which that index has been observed. If time is ignored or irrelevant, the exponent is 0. If the duration, or *quantity*, of time is measured, the exponent is  $+1$ ; while if something is expressed in units of duration, as in the velocity ratio, the exponent is  $-1$ . If, more penetratingly, the *relation* between the velocities of sequential periods is observed, the exponent is  $-2$ ; while if the *relation* between the durations of different kinds of things is observed, the exponent is  $+2$ .

### B. The Temporal Descripts, $\frac{t}{t}$

The thing whose duration or velocity is stated may be an index of any sector. Thus, one speaks of ages of people, of land, or of some characteristic, and of the velocity of processes such as mortality, acculturating, raising the marriage age, or expanding of territory— $tT^{+1} : (P, L^2, I)$  and  $(\%P, I_i, T_i^{+1} L^2) T^{-1}$ , respectively. Pure time may be multiplied by, or divided into, any index, creating indices of innumerable kinds. But pure time, a duration uncharacterized as to what it is a duration of, is one and but one homosectoral index. As such it can never have a plural class script,  $|_t$ , but always has the singular class script,  $|$ , whether explicit or implicit. However, when the duration is of some qualitative characteristic, there is an attribute-time product which is usually written implicitly by condensing the scripts of the attribute onto the time index:

$I^0 T_0^{+1} = T_i^{+1}$  = the implicit attribute-time product, duration of some attribute (Eq. 8a, Ch. IX)<sup>5\*</sup>

Thus, in S. 9, Ch. XII, the ages of husbands ( $I^0 T^{+1} = T_i^{+1}$ ) are cross-classified with the ages of wives ( $I^0 T^{+1} = T_{i'}^{+1}$ ). This attribute-time product is exactly parallel to the attribute-indicant product (Eq. 9, Ch. III), the attribute-population product (Eq. 4, Ch. IV), and the attribute-space product (Eq. 5, Ch. VIII) discussed in previous chapters. All these attribute-index products are usually written implicitly for economy, but must be written explicitly whenever the attribute ( $I^0$ ) is observed in degrees and becomes an indicant ( $I^{+1}$ ).

The great majority of situations have but one time index. The data can usually be graphed with time represented by one co-

\* For Eq. 8b, Ch. IX, see notes at end of chapter.

ordinate. (The chief exception is where ages are given for a series of calendar years, i.e., a situation of changing durations,  $\mathfrak{T}^{-1} : \mathfrak{T}^{+1} : .$ ) This one time index should not be confused with time measured in differing sets of units, such as years, months, reigns of kings, which are denoted by the period, or class-interval, script.<sup>6</sup>

The usual connotation here is that the periods are consecutive and of equal length. When the periods are not consecutive, or are of unequal length, this fact may be symbolized by the use of plural limits in the date or case script, as in Eq. 4d, Ch. IX above. Cycles, such as generations, business cycles, the rise and decay of institutions, are but one type of somewhat variable time periods, and multiples of them may be denoted by the period script. (See S. 2, Ch. IX and S. 7, Ch. XII.)<sup>7 \*</sup>

The date script, as the case script may be called on the time index, denotes a point in time.<sup>8 †</sup> It is useful, especially in specifying the limits of the period recorded in an S-situation, or in specifying the origin for reckoning durations, such as Anno Domini, birth-days for reckoning ages of people, or in specifying the present as a sliding zero point dividing the past from the future.<sup>9</sup>

## II. DURATIONS OF INDICES

### A. *Populational Durations, $T : P$*

Quantitatively recorded situations, in which the emphasis is upon the duration of something in time, are of less frequent occurrence than situations in which the emphasis is upon the change, whether slow or fast, of something in time. The most common of the durational situations are the durations of a population of which population pyramids are a large subvariety. The population pyramid is a distribution of people by ages, repeated for each sex. Its typical S-formula is:

$$\mathfrak{T}^{+1} : P_p = \text{a population pyramid} \quad (\text{Eq. 12, Ch. IX})$$

where  $\mathfrak{T}$  is the number of age class-intervals and  $p$  are the two sex plurals.

With this as a basis a large range of demographic facts and principles can be expressed in mathematical form. The tendency

\* For Eqs. 9a-b, Ch. IX, see notes at end of chapter.

† For Eqs. 10-11e, Ch. IX, see notes at end of chapter.

of a population pyramid to be normally isosceles is measurable by its negative correlation between age and population frequency:

$$T \bullet P < 0 \quad (\text{Eq. 13a, Ch. IX})$$

which is perfect (i.e.,  $T \bullet P = -1.00$ ) when the pyramid is isosceles. The degree of abnormality thus defined is measured by the discrepancy of the correlation coefficient from unity. (See S. 26, Ch. II.) For another example, the generalization that the sex ratio of males to females tends to vary inversely with the duration of life, can be expressed by an equation of the form:

$$T \bullet (735 - 700 P_{\text{group}}) > 0 \quad \text{rough relation of sex-ratio to age group}^{10} \quad (\text{Eq. 13b, Ch. IX})$$

This crudely represents the relation by a straight line from which the sex ratio for any given age may be predicted. More accurate prediction may be made from a fitted curve.

This illustrates the possibility of dealing mathematically with the many facts of societal significance which can be derived from the population pyramid. Among such are:

frontier plurels, indicated by excess of adult males

colonizing mother countries, indicated by excess of adult females

unhealthy plurels, indicated by concave pyramids (i.e., high infant mortality)

healthy plurels, indicated by convex ("beehive") pyramids. (Study S. 77, Ch. X.)

non-reproducing plurels, indicated by pinched-in bottoms, such as cities, or countries with falling birth rates

non-residential plurels, such as business groups, indicated by excess of adults. (See S. 26, Ch. II.)

plurels with major wars in the present generation indicated by indentation of males and babies of the war period from 1914-19. (Study France and Germany in S. 5, Ch. IX.)

stabilization of population composition, indicated by "waves" of bulges in the pyramid at different dates. (See S. 14, Ch. XII.)

migrational plurels as the foreign born in the U.S.. indicated by "bulges" in adulthood. (See S. 75, Ch. X.)

marital mores, indicated by dissimilar curves for sexes, and races as in S. 6, Ch. XII.

increasingly elderly populations, indicated by pyramids fatter at the top. (Study S. 76, Ch. X.)

and many other demographic phenomena.

The phenomena thus graphed are steadily being reduced to more rigorous mathematical treatment. Thus, from smoothed or fitted curves of populational durations predictions of social and economic importance can be made, as in the labor turnover of a factory illustrated in S. 7, Ch. IX. The largest field of these populational durations is in life, and other forms of insurance. Actuarial mathematics has developed the science of compound probabilities to a very high point. Starting from simple tables of longevities and expectancies of life (as in S. 29, Ch. II; S. 80, 81, Ch. X; S. 42, Ch. XI; S. 8, 12, 13, Ch. XII) and mortality rates by age, insurance calculations range up to very involved calculations. The high development of actuarial work and the enormous utility of insurance in the modern world are outstanding answers to those critics who assert that the social sciences cannot attain mathematical precision in their generalizations.

While critics outside the field of Sociology often ignore the achievements in the field of insurance, this field is also too often ignored by sociologists. The elementary textbooks of Sociology hardly refer to this field of practical achievement. The theory on which it is based is hardly emphasized, namely, the constancy of statistics from large samples or, more broadly, the laws of sampling. The principle, in measured form, is that the standard error of any index, in general, varies inversely as the square root of the size of the sample:

$$\sigma(I) \bullet (P^{-\frac{1}{2}}) = -1.0 = \begin{array}{l} \text{general relation of sampling error to size} \\ \text{of sample} \end{array} \quad (\text{Eq. 14a, Ch. IX})$$

This principle that the reliability of quantitative data tends to increase as the square root of the size of the population observed is relegated to statistical textbooks and hardly mentioned in current sociological textbooks. Yet it embodies one of the most

essential principles in Sociology on which rests, in large part, the hope of founding a science, of establishing uniformities.<sup>11</sup> \* Valid societal generalizations or laws are valid in large part because human phenomena in the mass or on the average are more stable and predictable than in the individual person. This principle of increasing stability of behavior with increasing size of sample observed, is almost as basic a working tool in the social sciences as the principle of conservation of energy is in Physics. Progressively as more sociologists realize this principle and utilize it as insurance actuaries do, Sociology may be expected to become more of an exact science and less of a literary discipline. Towards this objective a contribution of S-theory may prove to be its standardized symbols and matrix formulations, which should tend to bridge the existing gap between partially qualitative and crudely tabulated data on the one hand, and on the other hand, equations of best fitting curves or other calculative formulae which increase man's ability for prediction and controlled manipulation of societal phenomena.

### *B. Indicatory Durations, $T : I$*

After the class of situations involving durations of population, the class of durations next in importance is that of durations of characteristics. The duration of culture traits and complexes would be subsumed under this class, as also would much of the phenomena going under the conventional rubrics of "tradition," "custom," and "institutions" whenever their properties of persistence and continuity are being studied. Culture traits and complexes have longevities. They are youthful upstarts, or hoary elders surviving for centuries. The symbolic description for this class of societal phenomena, by the rules of S-notation, is:

$${}_tT^{+1} : I_{i,j}^1 = \text{longevity of culture traits and complexes} \\ (\text{Eq. 15, Ch. IX})$$

This denotes an age distribution of qualitative and quantitative ( $I^i = I^0, I^{+1}$ ) characteristics and their subclasses. When these characteristics have been created by man, their total aggregation is called culture, which, in turn, is subclassified into culture complexes in a hierarchy descending to the more-or-less elementary

\* For Eq. 14b, Ch. IX, see notes at end of chapter.

culture trait. The number of degrees (or levels) of subclasses in the hierarchy in any recorded situation is denoted by the number of class scripts,  $|_{ss} = |_{i+j+k+l+\text{etc.}}$ . Strictly speaking, Eq. 15 asserts durations of any characteristics, physical as well as cultural, but the cultural characteristics are the sociologically more important type of I in Eq. 15. When the indicator is a dynamic one ( $IT^{-1}$ ), Eq. 15 can symbolize customs and folkways, the enduring or repetitive activities of people.

The function of such a formula as this is partly to make the definition of the phenomena studied more precise and objective. The ages must be stated and not left indefinite in relative adjectives such as "very old," "quite new," etc., if  $\frac{1}{2}T$  is to be specified in the legend. The culture traits and complexes must be defined by some record, such as a schedule card with operational instructions for filling it out. Each item of the schedule card defines a trait, and any specified set of items defines a complex. These qualitative sets of data, or quantitative sum of weighted scores, define the indicators in the hierarchy of culture,  $I_{i,j}^1$ .<sup>12</sup>

Proportionately as cultural anthropologists define the somewhat loose and unstandardized "culture traits" and "culture complexes" in operational terms of exactly specified questions to be asked, and observations and distinctions to be made in filling out a schedule card, greater objectivity and precision of the facts of culture will result.

A second function of such a formula as Eq. 15, Ch. IX is, in common with all descriptive S-formulae, to serve as a base for manipulating and studying the relationships in the situation. Comparisons, or equalities and inequalities, correlations and contingencies, sequences, classificational patterns, and many other relationships within the situation, can then be readily and precisely denoted, either as facts, or as hypotheses to be tested by patterns (i.e., variant formulae) built up of the entities defined by the legend of Eq. 15. Thus, in S. 22, Ch. VIII the longevities of the culture complex in the countries of Europe, entitled "national independence," are summarized in a "national index" and correlated with indicators of language, rainfall, religion, health, race, and illiteracy. From Eq. 12, a population pyramid, a dozen relations such as Eq. 13, the relation of sex-ratio to age, can be built up as listed in the paragraph following Eq. 13. Hitherto,

any such relation as Eq. 13 has had to have its symbols defined *ad hoc* for that situation, but the standardized notation of S-theory both fits that situation into a class in a comprehensive system of classification, and facilitates deriving further relations by merely varying, in any recorded situation that is being studied, the permutations and combinations of the entities defined by the legend of Eq. 12.

For another example, whereby the formulation in algebraic notation can sharpen and order a temporal sociological classification consider Lundberg's methodological classification of sociological observation by sources (Ref. 45, p. 86) as:

- I. Historical sources
  - a. Geological strata
  - b. Documents, inscriptions, etc.
- II. Field sources
  - c. Information from living individuals
  - d. Direct observation of behavior

These are an aggregation of qualitative characteristics  $I^0_{i,j}$ , where  $i = 2$  and  $j = 2$ . But they form a time series, correlating highly with successively receding and lengthening periods in the past which can be expressed by the successive exponents on a suitable time period which is here a century ( $T = 100$  years). The formula for this series of periods then is:

$T^t : I^0$  = a logarithmic series of qualitatively characterized periods (Eq. 16, Ch. IX)

This series can be extended by increasing the exponent beyond +3, to divide the past into equal logarithmic periods which, as a mnemonic device for cosmic development, may be qualitatively characterized approximately as tabulated below:

$t$		$T^t$	$I^0$
0	a year	$= 100^0$	the <i>societal present</i> , II <sub>d</sub> above.
1	a century	$= 100^1$	the maximal period within which <i>persons</i> now living have developed, II <sub>c</sub> above.
2	10 millennia	$= 100^2$	the maximal period within which <i>civilization</i> has developed, i.e., recorded, history, I <sub>b</sub> above.

- |   |                   |           |   |
|---|-------------------|-----------|---|
| 3 | 1 million years   | = $100^3$ | the maximal period within which <i>man</i> has developed, Ia above.   |
| 4 | 100 million years | = $100^4$ | the probable period within which <i>life</i> has developed on our planet.   |
| 5 | 10 billion years  | = $100^5$ | the maximum period within which our <i>solar system</i> has developed including the birth of our sun and its planets. |

By shifting the time base from a century to a decade ( $T = 10$ ), intermediate periods may be marked off, such as periods in which modern civilization, stone age man, mammals, etc., have developed. This extends the series from six to eleven items as now  $T^6 = 10^{10}$  at maximum.

### C. Spatial Durations, $T : L^1$

A final type of durations is spatial durations. Although important in Geology, these are of less importance in Sociology. This includes the distributions by longevity, or by duration up to the present, of spatial points, lines, areas, or volumes of specified types. This formula is:

$$\{T^{+1} : \{L^1\} = \text{spatial durations} \quad (\text{Eq. 17, Ch. IX})$$

Data of this sort occur in studies of title deeds to land; necessities of crop rotation; wear of roadbeds, cables, pipes; age of city and archaeological ruins; age of containers, from freight cars to cranial capacities, etc. Here again, the descriptive formula is a basis for expressing relationships in the situation. Thus, in S. 12, Ch. IX, the cranial capacities of persons of ages 6 to 19 can be compared by sex, stating the principle of a smaller capacity of females at all ages, in a rough equation of the form:

$$\underline{P}_r : \{T^{+1} : L^3 = \underline{P}_r : \{T^{+1} : L^3 + 60 \pm \quad (\text{Eq. 18a, Ch. IX})$$

This states that for males of ages 8 to 19 the cranial capacity is approximately 60 cc. greater than that of females of those ages. Given the legend of S. 12, Ch. IX, Eq. 17 states, to one who is practiced in verbalizing the symbols of S-theory, a generalization from the data without the need of any further legend, such as the

previous sentence. (For a briefer equation see Eq. 18c of this chapter.)

Again the mean amount of increase of cranial capacity per year in this age range is, for males:

$$\underline{P}_r : L^3 T^{-1} = 17.7 \quad (\text{Eq. 18b, Ch. IX})$$

For another example of manipulating the basic descriptive formula of a situation, consider the hypothesis that the smaller cranial capacity of females is proportionate to their smaller bodies. Denoting the new variable, "body size," measured by "cubic centimeters of water displaced on immersing" as  $L^3$ , the hypothesis is stated as:

$$\underline{P}_r : L^3_{f,m} ? = \underline{P}_{f,m} : L^3_{f,m} \quad (\text{Eq. 19a, Ch. IX})$$

For those not practiced in verbalizing S-formulae, this states that, for males, the ratio of cranial to body capacity is, according to hypothesis, equal to that of females. (The question mark denotes a question as to the equality, an uncertainty to be tested. It labels the equation not as a fact, but as an hypothesis.) An hypothesis such as this has obvious significance for the issue of the equality of the sexes. The hypothesis tends to be true, though the data for proof are not provided in S. 12, Ch. IX as recorded. Many further hypotheses of relation can be explored, such as whether this cranial capacity corrected for body capacity has more, or less, correlation with intelligence than the cranial capacity alone (which has little correlation to intelligence within the human range, but has higher correlation in the larger anthropoid range).

Equations 17, 18, and 19 illustrate a function of S-formulae in aiding the exploration of relations in a given situation (such as S. 12, Ch. IX, which is a particular application of the more general type of spatial durations, Eq. 16, Ch. IX). Of course, these equations neither discover nor verify such relations. They simply record, in a form suitable for mathematical reasoning, the insights and findings of the researcher (with whatever degree of precision he has achieved in his observing), in a standardized notation which then automatically classifies these data into their quantic class in the whole field of societal phenomena.

## III. BRIEF S-FORMULA

In the preceding chapters up to this Part V, the notation for the four sectors has been cumulatively expounded.<sup>13</sup> The remaining chapters add no new symbols. Consequently, the notational system may be briefly reviewed at this point and a condensed or briefer notation pointed out.

All the full descriptive or matrix formulae up to this point in the exposition have excluded cross-scripting. Cross-scripting means the attaching of scripts from one sector onto a base letter from another sector. In the full descriptive formula the scripts are always homosectoral, i.e., of the same sector and denoted by the same letter, as the base letter to which the descripts may be attached. But when cross-scripting is permitted, a flexible, compact, but less precise notation, designated "Brief-S formulae," results. As these Brief-S formulae will be much used in the remaining chapters, they may be explained herewith.

The aim of Brief-S formulae is parsimony of symbolism. Its technic is to select the rightmost index, i.e., the most dependent index in an S-situation, and to denote what the other indices denote by attaching their scripts to it. This is "cross-scripting." Thus, a frequency distribution of people on some characteristic is by full formula:

$$\begin{aligned} &{}_i\mathfrak{I} : \mathfrak{P}, & (\text{Eq. 20, Ch. IX}) \\ \text{or } &{}_i\mathfrak{P} \text{ by brief formula} \end{aligned}$$

This states an aggregation of  $\mathfrak{P}$  persons at each of  $i$  class-intervals of an indicator,  $\mathfrak{P}$  being a variable with  ${}_i|$  values.<sup>14</sup> \*

The convenience of these brief formulae is often very great, so that they will be used extensively in defining societal processes in the chapters ahead. For examples of their brevity, note how the three equations summarizing the facts, or stating an hypothesis, with regard to the spatial durational situation of cranial capacity of the sexes increasing with age as graphed in S. 12, Ch. IX, can be condensed:

$$\begin{aligned} &{}_t\mathfrak{L}^3 \bullet ({}_t\mathfrak{L}^3, + 60) = r > 0 & (\text{Eq. 18c, Ch. IX}) \\ & & (\text{Cf. with Eq. 18a}) \end{aligned}$$

\* For Eqs. 21-28, Ch. IX, see notes at end of chapter.

This states that the volume ( $L^3$ , cranial capacity) for  $t$  periods (ages) of males,  $|_t$ , tends to equal the volume for those ages of females increased by 60 cc. The tendency to equal is denoted by the correlation symbol  $\bullet$ . This usage should be noted carefully, for it is of great import for the social sciences. Exact equations between observed societal indices are seldom possible, but their correlations can be observed. All correlations of indices mean a tendency to equal each other (when expressed in sigma units). When the correlation is 1, the tendency equation or regression equation becomes the exact equation of mathematics. When  $r$  is less than 1, as it usually is, the equation, properly manipulated, can still yield much of the service that equations yield in Chemistry and other sciences.

The velocity of growth of cranial capacity simplifies to:

$$\underline{P}_t : L^3 T^{-1} = 17.7 = L^3_{/t} \quad (\text{Eq. 18d, Ch. IX})$$

Here the "Brief-S" at the right asserts a volume per period of a particular plurel being equal to 17.7 cc. Again, the equality of the sexes in regard to the ratios of cranium-to-body capacity in Eq. 19a, Ch. IX becomes:

$$L^3_{///(p? = p'')} \quad (\text{Eq. 19b, Ch. IX})$$

From the legend for S. 12, Ch. IX, Eq. 19b can be directly verbalized as: the ratio of cranial to body,  $|_{///}$ , capacity,  $L^3$ , of males,  $|_{p'}$ , equals, according to hypothesis,  $? =$ , that ratio of females,  $|_{p''}$ . For further examples of Brief-S formulae, see the standardized patterns of S-formula in the Appendices.<sup>15</sup>

#### IV. S-SITUATIONS

For examples of the durational formulae of this chapter the sample of twelve S-situations which follow may be studied. These furnish a little evidence of the extent to which the S-symbols fit the observed and currently published quantified data in the social sciences and bring them all within the S-classification of which the major classes are diagrammed in S. 33, Ch. II. These S-situations are a small sample of the fifteen hundred photostated sets of data culled from the literature of the social sciences of the past score of years from which the S-symbols and formulae were induced.

## S. 1

## THE WORLD CALENDAR

All Years Alike  
All Quarters Equal

First Quarter							Second Quarter							Third Quarter							Fourth Quarter						
JANUARY							APRIL							JULY							OCTOBER						
S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S
1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
8	9	10	11	12	13	14	8	9	10	11	12	13	14	8	9	10	11	12	13	14	8	9	10	11	12	13	14
15	16	17	18	19	20	21	15	16	17	18	19	20	21	15	16	17	18	19	20	21	15	16	17	18	19	20	21
22	23	24	25	26	27	28	22	23	24	25	26	27	28	22	23	24	25	26	27	28	22	23	24	25	26	27	28
29	30	31	...	...	...	...	29	30	31	...	...	...	...	29	30	31	...	...	...	...	29	30	31	...	...	...	...
FEBRUARY							MAY							AUGUST							NOVEMBER						
S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S
...	...	1	2	3	4	...	...	...	1	2	3	4	...	...	...	1	2	3	4	...	...	...	1	2	3	4	...
5	6	7	8	9	10	11	5	6	7	8	9	10	11	5	6	7	8	9	10	11	5	6	7	8	9	10	11
12	13	14	15	16	17	18	12	13	14	15	16	17	18	12	13	14	15	16	17	18	12	13	14	15	16	17	18
19	20	21	22	23	24	25	19	20	21	22	23	24	25	19	20	21	22	23	24	25	19	20	21	22	23	24	25
26	27	28	29	30	...	...	26	27	28	29	30	...	...	26	27	28	29	30	...	...	26	27	28	29	30	...	...
MARCH							JUNE							SEPTEMBER							DECEMBER						
S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S
...	...	...	...	...	1	2	...	...	...	...	...	1	2	...	...	...	...	...	1	2	...	...	...	...	...	1	2
3	4	5	6	7	8	9	3	4	5	6	7	8	9	3	4	5	6	7	8	9	3	4	5	6	7	8	9
10	11	12	13	14	15	16	10	11	12	13	14	15	16	10	11	12	13	14	15	16	10	11	12	13	14	15	16
17	18	19	20	21	22	23	17	18	19	20	21	22	23	17	18	19	20	21	22	23	17	18	19	20	21	22	23
24	25	26	27	28	29	30	24	25	26	27	28	29	30	24	25	26	27	28	29	30	24	25	26	27	28	29	30

\*YEAR-END DAY, December Y, follows December 30th every year

\*\* LEAP-YEAR DAY, June L, follows June 30th in leap years

The World Calendar is a revision of the present calendar to correct its inequalities and discrepancies. It rearranges the length of the 12 months so that they are regular, making the year divisible into equal halves and quarters in a "perpetual" calendar. Every year is the same; every quarter identical.

In this new calendar, each quarter contains exactly three months, 13 weeks, 91 days. Each quarter begins on Sunday and ends on Saturday. The first month in each quarter has 31 days, and the other two 30 days each. Every month has 26 weekdays.

In order to make the calendar perpetual (identical for every year), at the same time retaining astronomical accuracy, the 365th day of the year, called Year-End Day, is an intercalary day placed between December 30th and January 1st, and considered an extra Saturday. The 366th day in

leap years, called Leap-Year Day, is intercalated between June 30th and July 1st on another extra Saturday. These intercalary or stabilizing days are tabulated as December Y and June L, and would probably be observed as international holidays. January 1st, New Year's Day, always falls on Sunday.

The revised calendar is balanced in structure, perpetual in form, harmonious in arrangement. It conforms to the solar year of 365.2422 days and to the natural seasons. Besides its advantages in economy and efficiency, it facilitates statistical comparisons, co-ordinates the different time-periods, and stabilizes religious and secular holidays. As compared with any other proposal for calendar revision, it offers an adjustment in which the transition from the old to the new order can be made without disturbance.

"Our stability is but balance."—Robert Bridges.

*Descriptive formula:*  $S_1 = t : u : v : w : x T^{+1}$

*Quantic number* = 1;0;0;0

*Legend:*

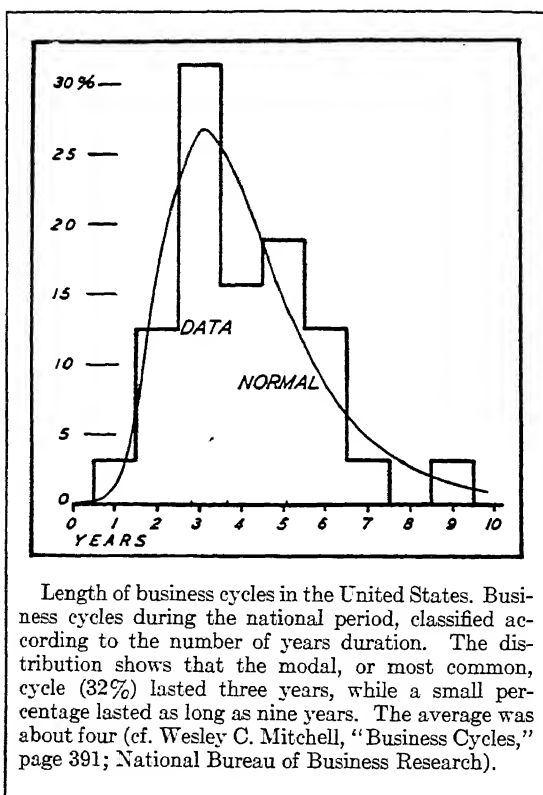
$S_1$ = The situation	$v $ = 3 second-order subperiods, the
records	months
$t $ = all periods	each
of	: = subclassified into
$T$ = a duration of one year	$u $ = 4 third-order subperiods, the
: = subclassified into	weeks
$u $ = 4 first-order subperiods, the seasons	: = each subclassified into
: = each subclassified into	$x $ = 7 fourth-order subperiods, the
	days.

*Comment on notation:*

1. This S-situation illustrates a hierarchy of period scripts which are *class-interval scripts* in the temporal sector  $T$ . Since the total period is classified in five different ways, on the basis of five different units of time (years, seasons, months, weeks, days), the formula could be written more compactly, but with loss of detail, as  $T^{+1}$ , denoting a duration of a year  $T$ , in five *classifications*,  $|t$ . There are thus five indices. In vectorial terms there are five congruent *vectors*; which can also be considered as one vector subdivided into five sets of line sects. (Cf. S. 35, Ch. II.)

2. The meaning of a *matrix*, an aggregation of numbers in a rectangular arrangement, is also illustrated. The total matrix is subdivisible into submatrices. Thus, a month is an imperfect  $7 \times 5$  submatrix, meaning that it has seven columns and five rows, but is imperfect in that it has some empty cells. A week is a  $7 \times 1$  submatrix of the month matrix. One day is the lowest limit of the hierarchy of matrices, as it is a matrix of only one cell, i.e., a  $1 \times 1$  matrix.

## S. 2



Ref.: Reinhardt and Davies, *Principles and Methods of Sociology*, Inc., 1932, p. 533.

Descriptive formula:  $S_2 = t : \sum_u T_t^{+1}$

Quantic number = 1;0;0;0

Legend:

$S_2$  = The situation

is a record of

$tT^{+1} = 10$  lengths of periods

each with a corresponding

$\sum_u$  = frequency of depressions

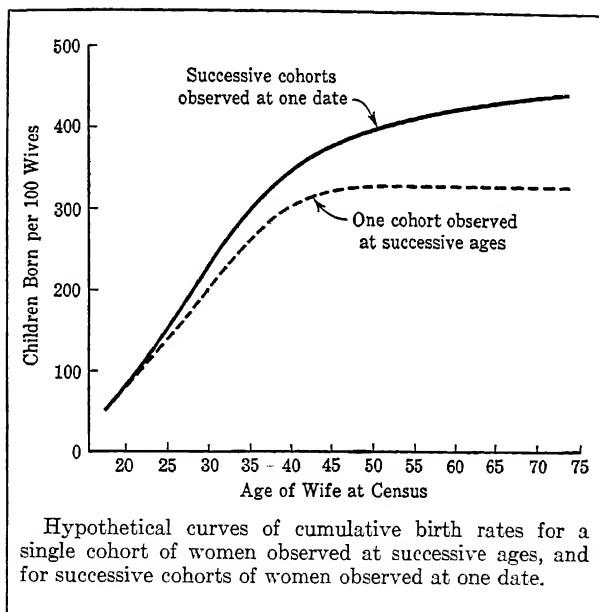
$t$  = in 2 classifications

{ observed  
distribution  
and  
normal  
distribution

Comment on notation:

The  $\Sigma$  symbol before a descript converts it from denoting an aggregation to denoting a single number. The total symbol  $S_2$  denotes a homosectoral frequency distribution in the time sector. It asserts an aggregation of periods,  $t$  in number, each with its corresponding ( $\sum_u$ ) frequency of occurrence.

## S. 3



Ref.: Sallume, X., and Notestein, F. W., "Trends in the Size of Families (computed prior to 1910) in Various Social Classes," *Amer. Journ. Soc.*, Vol. XXXVIII, No. 3, Nov., 1932, p. 401.

Descriptive formula:  $S_3 = ({}_tT^{+1,-1} : \Sigma P, P_{,,}^{-1})_t$

Quantic number = 19;0;0;19

Legend:

$S_3$  = The situation

records

$|_t$  = 2 classifications of time

into

${}_tT^{+1}$  = a series of durations (ages)

and

${}_tT^{-1}$  = growths (rates of change)

each classification having a corresponding

$\Sigma$  = cumulative

$P,$  = frequency of births

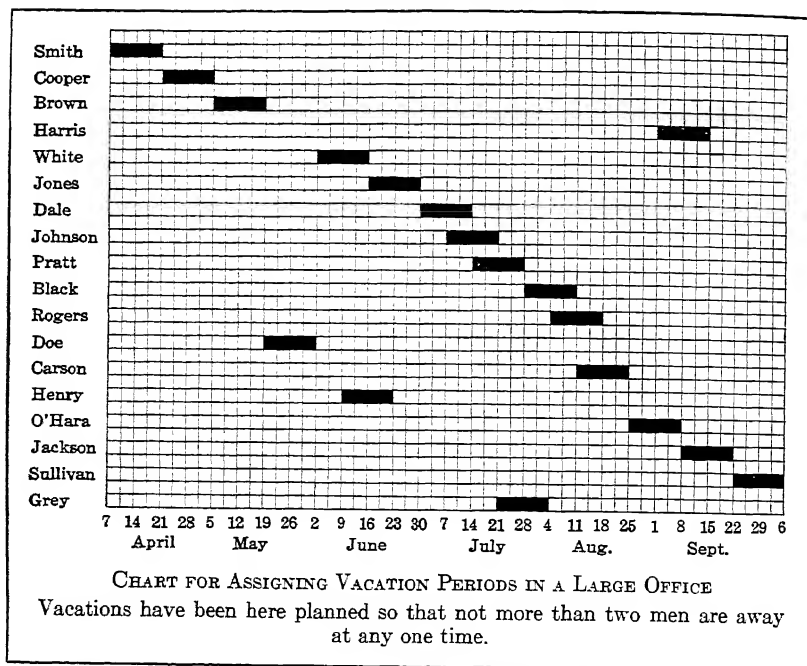
$P_{,,}$  = per 100 wives

Comment:

This situation illustrates time as duration and rate of change, symbolized in S-notation by exponents of plus one and minus one respectively. (See S-Rules #12-17 in Appendix II.)

## S. 4

## VACATIONS FOR THE YEAR 1912



Ref.: Brinton, Willard C., "Graphic Methods for Presenting Facts," Engineering Magazine Co., 1923, p. 53.

Descriptive formula:  $S_4 = {}^pP : {}^a : {}^{z'}T$

Legend:

$S_4$  = The situation  
records  
for each of

${}^pP$  = 18 employees

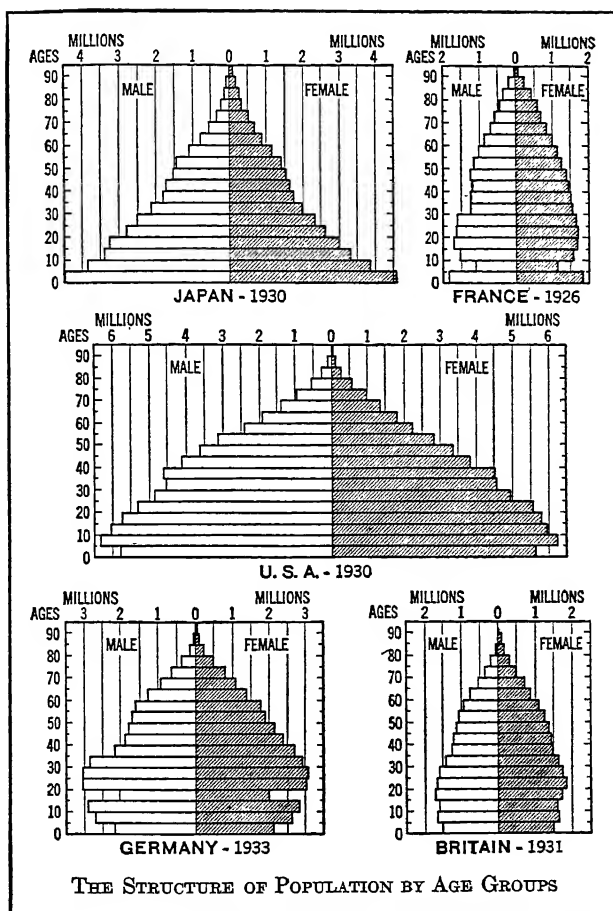
$:$  = a corresponding  
 $T$  = 2 weeks vacation  
with

${}^a : {}^{z'}$  = initial and terminal dates

Comment on notation:

This situation is a list distribution of periods by persons. (Cf Eq. 1b, Ch. V.) It illustrates a use of S-notation to describe calendrical events, either as a historical record of the past or as predictions in a plan for the future.

## S. 5



THE STRUCTURE OF POPULATION BY AGE GROUPS

Ref.: Yano, T., and Shirasaki, K., *Nippon, A Chartered Survey of Japan*, Kikusei-Sha, Tokyo, 1936, chart 171, p. 445.

Descriptive formula:  $S_s = ({}_tT^{+1} : P_p)_q : T^0$

Legend:

$S_s$  = The situation

records

${}_t|$  = 20 age periods

$T^{+1}$  = each of 5 years

each with a corresponding

P = population

Quantic number = 1;0;0;1

subdivided by

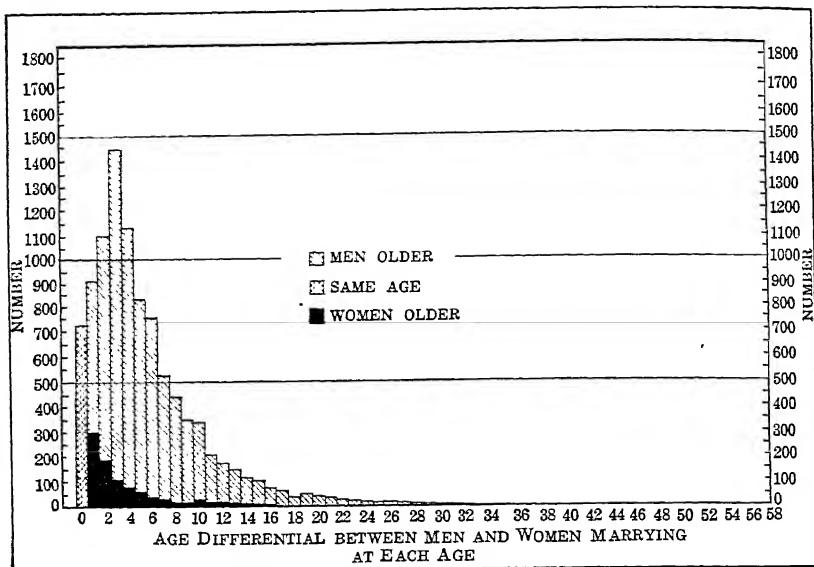
$|_p$  = the 2 sexes

and all repeated for each of

$|_q$  = 5 nations

$T^0$  = at a certain date for each nation

## S. 6



Ref: Duncan, Otis, "The Factor of Age in Marriage," *Amer. Jour. Soc.*, Vol. XXXIX, No. 4, Jan. 1934, p. 476.

Descriptive formula:  $S_6 = \iota T^{+1} : P$

Quantic number = 1;0;0;1

Legend:

$S_6$  = The situation  
records

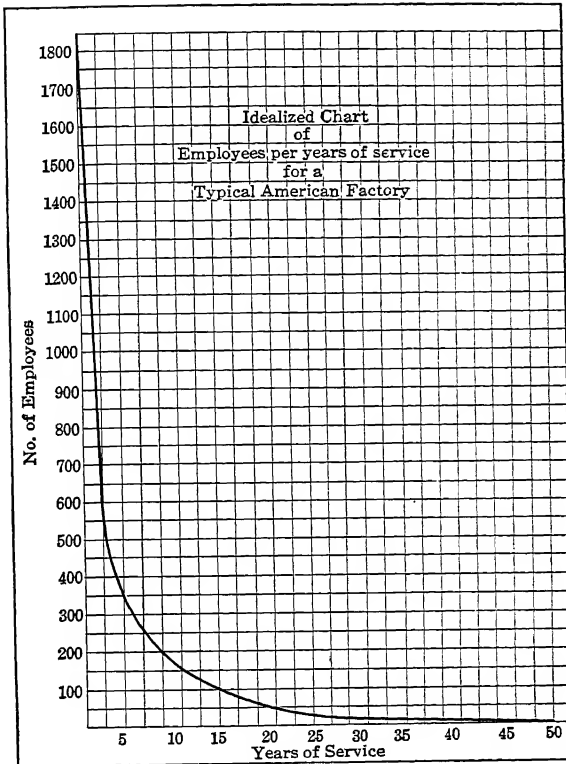
= with a corresponding  
P = frequency of husbands and  
wives

$\iota T$  = some 50 intervals of age dif-  
ference

Comment:

The probability of the age discrepancy exceeding any assigned amount,  $\iota T$ , is, using cross scripts,  $\iota P/P$ , the population exceeding a particular time interval divided by the whole population, P. More refined calculation of probabilities might be attained by fitting a normal distribution curve to these data.

S. 7



Ref.: James, Gorton, "Baby Annuities," *Survey*, Vol. XLVI, No. 20, Sept. 24, 1921, p. 707.

Descriptive formula:  $S_T = {}_tT^{+1} : P$

Quantic number = 1;0;0;1

Legend:

$S_T$  = The situation  
is a record of  
 ${}_t$  = 50 periods  
of

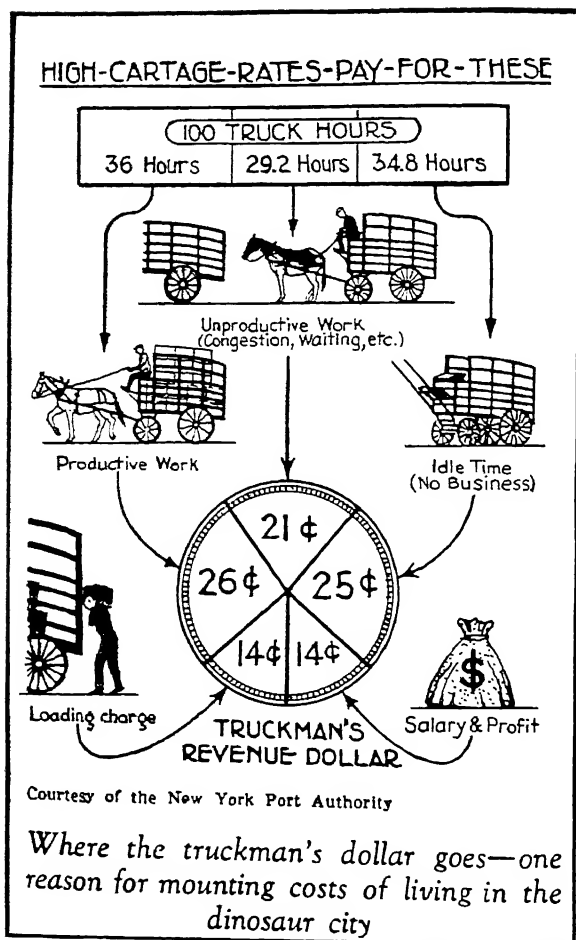
$T$  = a year's service  
: = each with a corresponding  
 $P$  = number of employees

Comment:

Of course the equation of some best-fitting curve would define this curve with greater determinateness.

This S is a borderline one between duration and change,  $T^{+1}$  and  $T^{-1}$ . Duration dominates if the data are a survey on one date making a J-shaped distribution curve of employees by length of service. (See S-Rule #12 in Appendix II.) Change dominates if one starts with a standard employee population on a certain date and plot decrements of it every year thereafter, i.e., the number fired or the number surviving out of that original population.

## S. 8



Ref.: Stein, Clarence S., "The Dinosaur Cities," *Survey*, Vol. LIV, No. 3, May 1, 1925, p. 138.

Descriptive formula:  $S_s = {}_sI_1 : {}_sT^{+1}$

Legend:

$S_s$  = The situation

records

$I_1$  = 5 kinds of trucking cost

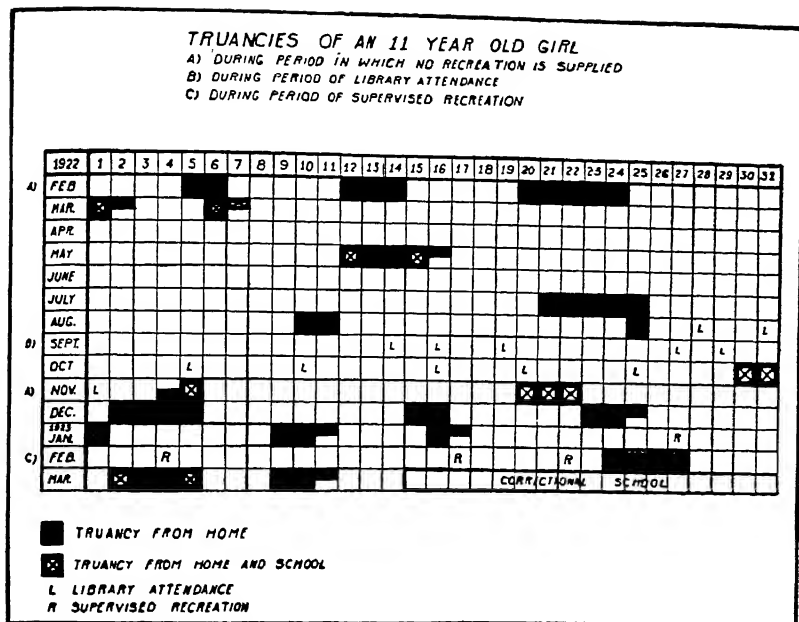
Quantic number = 1;1;0;0

with corresponding

$T^{+1}$  = time

${}_s|$  = all in percentage units

## S. 9



Ref.: Trunip, Elizabeth, "What Does the Social Worker Do?" *Social Forces*, Vol. III, No. 2, Jan., 1925, p. 273.

Descriptive formula:  $S_9 = 'P : t : u T^{+1} : {}^1 I_i$

Quantic number = 1;1;0;1

Legend:

$S_9$  = The situation

records for

'P = a particular girl

$t T^{+1}$  = 14 months

$u$  = and their days

: = for each of which there was

${}^1 I_i$  = an all-or-none indicant

$|_i$  = of 5 kinds—

{ presence at home,  
 school, library,  
 playground and  
 correctional school

Comment:

This is an example of an attempt at causal analysis of a case with graphical aid.

Correlations with time sequence noted are the necessary, but not necessarily sufficient conditions of causation.

## S. 10

The results of the first series of tests are given in Tables V to XVI inclusive. Following is the key to the contents of all the tables of Series I.

*Ind.* = Individual. The name is omitted for obvious reasons.

*Age.* = Age in years.

1 = Assenting to overstatement regarding his grades.

2 = Yielding to the suggestion that he made a mistake.

3 = Accepting help contrary to his instructions.

4 = Returning or delivering articles according to promise.

5 = Keeping ten cents overchange.

6 = Accepting ten cent tip for small favor.

7 = Pushing a button according to instructions.

8 = Counting A's in a picture book as rapidly as in an uninteresting book.

9 = Peeping at a profile when on honor to keep his eyes closed.

10 = Cheating by adding words after test was over.

*M* = Merits and demerits. Merits are indicated by the plus sign; demerits by the minus sign.

*G* = Grade. This is obtained by adding the scores and the merits and demerits.

*R* = Rank given by the tests.

*RL* = Rank given by the judgment of the teacher or leader.

*IQ* = Intelligence Quotient.

TABLE V  
SERIES I. GROUP A

<i>Ind.</i>	<i>Age</i>	1	2	3	4	5	6	7	8	9	10	<i>M</i>	<i>G</i>	<i>R</i>	<i>RL</i>	<i>IQ</i>
1	10	10	2	10	0	10	10	5	10	10	10		77	3	2	148
2*	10	0	3	10	10	0	0	2	7	0	0	-10	22	7	7	133
3	11	0	4	5	10	0	0	2	10	7	0		38	6	5	142
4	11	2	10	10	0	10	10	10	10	10	10		82	2	3	151
5†	13	10	9	10	10	10	0	6	7	0	0	-20	42	5	6	135
6††	10	10	10	10	10	0	10	8	10	0	10	-10	68	4	4	146
7	11	10	3	10	10	10	10	10	10	10	5		88	1	1	139

\* Removed money from a pocket book which he found and then returned the book to the Lost and Found Department.

† Told a falsehood on his own initiative; stopped in the midst of an assigned task to attend to other matters of interest to him.

†† Stopped in the midst of an assigned task to attend to other matters of interest to him.

Correlation between the ranks obtained by the tests and those given by the four teachers = .928.

Ref.: Voelker, Paul Frederick, *The Function of Ideals and Attitudes in Social Education*, Teachers College, Columbia Univ., N. Y., 1921, p. 93.

Descriptive formula:  $S_{10} = {}^pP : (I_i, T^{+1})$

Quantic number = 1;1;0;1

Legend:

$S_{10}$  = The situation

$I_i$  = 15 indicants of character and other tests

records

$T^{+1}$  = and also their ages

for each of

${}^pP$  = 7 persons

*Comment on notation:*

The intersectoral structure of this situation is a "loose" one in contrast to one subclassifying indices of one sector by indices of another sector, or in further contrast to the "tight" structure of a cross-classification. The  $T$  and the  $I$ 's here are merely compared, as shown by the comma, and are bound together by their common dependence on the persons observed, as symbolized by the parenthesis on all of whose contents the colon and its preceding descripts operate impartially.

## S. 11

## DEGREE OF INTIMACY (SPRING 1932 GROUP)

<i>Pairs</i>	<i>No. of Cases</i>	<i>Mean Age Difference</i>	<i>Mean Scale Position</i>	<i>Mean Distance</i>	<i>Index of Intimacy</i>
(1)	(2)	(3)	(4)	(5)	(6) Based on Col. 4
Brothers.....	79	3.88	6.12	.97	117
Sisters.....	52	5.11	5.20	1.11	100
Totals or Means	131	4.37	5.75	1.02	108
Boy-Cousin.....	24	2.28	6.42	1.56	123
Girl-Cousin.....	17	4.31	4.69	1.47	90
Totals or Means..	41	3.13	5.70	1.52	106
Boy-Uncle.....	23	21.43	7.47	1.70	143
Girl-Aunt.....	16	21.97	6.61	1.91	127
Totals or Means..	39	21.65	7.12	1.79	135
Sisters.....	52	5.11	5.20	1.11	100
Girl-Cousin.....	17	4.31	4.69	1.47	90
Girl-Aunt.....	16	21.97	6.61	1.91	127
Means.....	85	8.12	5.36	1.33	105
Brothers.....	79	3.88	6.12	.97	117
Boy-Cousin.....	24	2.28	6.42	1.56	123
Boy-Uncle.....	23	21.43	7.47	1.70	143
Means.....	126	6.78	6.42	1.21	127

Ref.: Chapin, F. Stuart, *Contemporary American Institutions*, Harper and Brothers, 1935, p. 402.

*Descriptive formula*:  $S_{11} = P_{p,\Sigma p} :: P_q : (I_i, T)$       *Quantic number* = 1;1;0;2

*Legend*:

$S_{11}$  = The situation

records

P = a number of responders

in each of

$p,\Sigma p$  = 2 sex plurels, and their combination

who show towards

P = relatives of the same sex

classified in

$|_q$  = 3 kinship  
plurels

{ sibling  
cousin  
aunt-or-  
uncle

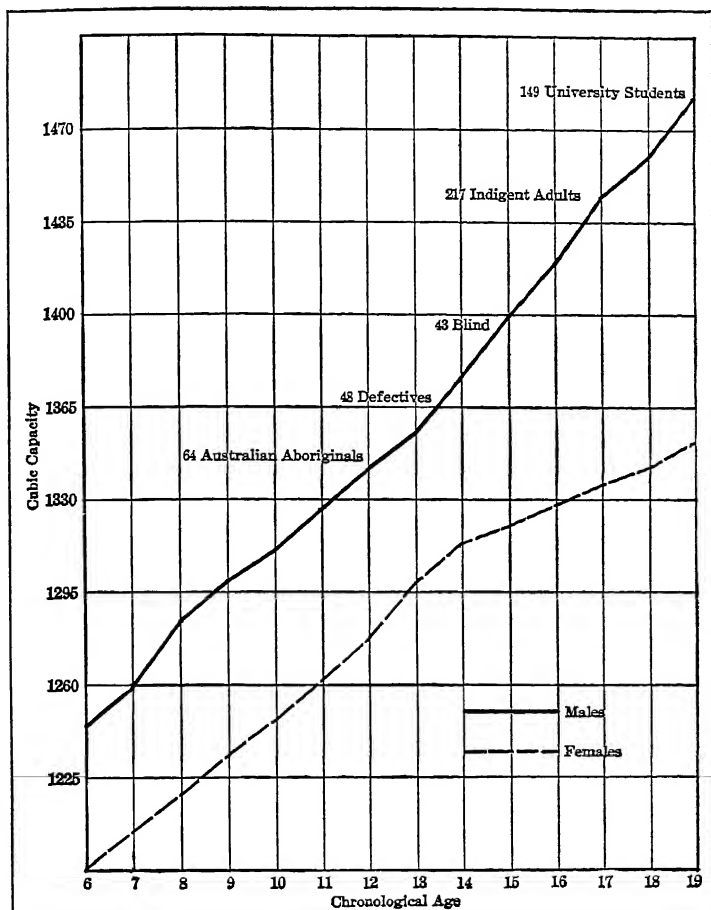
$I_i$  = 3 indicants of intimacy  
and

$,T$  = an age discrepancy is also  
stated.

*Comment*:

Interrelations being, by definition in S-theory, a stimulus-response type, it is the intimacy indices and not the age difference that calls for an analysis into cross-classified plurels with a quantic of  $P^2$ . Compare with S. 6, Ch. IX, where the age differences between husbands and wives were analyzed as a distribution of a temporal characteristic and not an interrelation of parties.

## S. 12



This graph illustrates two researches. It shows, first, the comparative growth of brain size in cubic centimeters for the two sexes, as shown by measurements on several thousand primary, secondary, and university students. Secondly, it shows comparative brain development in Australian Aborigines, three classes of defectives, and university men. Figures in all cases are only approximate, since measurements on living persons are subject to a considerable margin of error. From S. D. Porteus and M. E. Babcock, *Temperament and Race*, Richard G. Badger, 1926, by permission.

*Descriptive formula:*  $S_{12} = \underline{P}_{p,q} : \frac{1}{2}T^{+1} : L^3$

*Quantic number* = 1;0;3;1

*Legend:*

$S_{12}$  = The situation

$|_q$  = 5 special plurals

records

for all

for each of

$\frac{1}{2}T^{+1}$  = ages of life

$L^3$  = the cc. of brain capacity

$\underline{P}_p$  = 2 sexes,

, = with which are compared

## V. NOTES

1. As usual, an index with a zero exponent and a zero descript does not affect the situation and is called a nul index.

$T_0^0 = 1$  = the nul temporal index,  $|^s = 0; i; l; p$  (Eq. 1a, Ch. IX)

For a situation on a particular date, i.e., at a specified point of time, the singular date script is used:

$T^0$  = a specified date,  $|^s = 0; i; l; p$  (Eq. 1b, Ch. IX)

A situation involving an aggregation of dates which are not in sequence is symbolized by the plural date script and zero exponent:

$T^0$  = an aggregation of dates not in sequence,  $|^s = 0; i; l; p$  (Eq. 1c, Ch. IX)

Since the class-interval script denotes a period or an extension of time, it is incompatible with a zero exponent.

2. The standardized symbols combining these exponents and the common descripts may be noted by the advanced student who desires a mastery of the S-symbolism, sufficient for writing formulae for himself. These footnotes may be ignored by the reader who simply desires a recognition knowledge to follow analyses which are verbalized in their accompanying legends.

$\frac{1}{2}T^{+1}$  = a specified duration, a particular named period (Eq. 2b, Ch. IX)

$\frac{1}{2}T^{+1}$  = a series of successive durations, such as ages (Eq. 2c, Ch. IX)

$T^{\frac{1}{2}}_i$  = a duration, or age, of some specified kind of thing (this descript is understood to be present, even if it is not written) (Eq. 2d, Ch. IX)

=  $I^0T^{\frac{1}{2}}_0$  = the implicit attribute-time product, a qualitative characteristic times a pure duration of it (Eq. 2e, Ch. IX)

$\frac{1}{2}T^{+1}$  = a duration between a specified initial and terminal date (Eq. 2f, Ch. IX)

$(IT^{-1})$  = the velocity of change in I in a specified or named period,  $|$ . I is a dynamic index, i.e., acts, events, happenings over a period (including income and expense) (Eq. 3b, Ch. IX)

$\frac{1}{2}(IT^{-1})$  = the velocities of change in I in each of  $\frac{1}{2}$  periods, the amount of I per period for  $\frac{1}{2}$  periods (Eq. 3c, Ch. IX)

$T^{-1} : I$  = the velocities of change in a non-dynamic index, I, between successive dates,  $\frac{1}{2}$  in number (Eq. 3d, Ch. IX)

$\iota T^{-1} : I$  = the velocities, for each of  $\iota$  periods, of change in a non-dynamic index. This states for an aggregation of  $\iota$  successive periods, the amount of the change in the index,  $I$ , per period,  $T^{-1}$  (Eq. 3e, Ch. IX)

$\iota T^{-1} : I = \iota T^{-1} I$  = a velocity in a specified period. The colon becomes unnecessary in the singular case. It denotes in an aggregation, or a series, that for each period there is a corresponding amount of change. But for one period, "a change corresponding to one period" is the same as "a change in one period," or as "a change per that period" (Eq. 3f, Ch. IX)

3. Again the advanced student should note the commoner patterns of acceleration formulae which result from the standard meanings of the four scripts in their singular and plural forms, with the operational symbols relating them together.

$\iota T^{-1} : (IT^{-1})$  = accelerations in each of  $\iota$  successive periods in a dynamic index (events per unit of time) (Eq. 4b, Ch. IX)

$\iota T^{-1} : \iota T^{-1} : I$  = the acceleration in a particular period,  $\iota$ , of the velocities of change in  $\iota$  subperiods of a static index (amounts on a given date, i.e., observable at an instant of time, not requiring a period in which to be cumulated) (Eq. 4c, Ch. IX)

The chief complexities that are likely to occur are in combinations of accelerations with durations of some per capita index and when, further, the time periods are irregular:

$\iota : \iota' T^{-1} : \iota : \iota' T^{-1} : (\%PT^{-1})_p$  = accelerations of a population rate distributed by age in several plurals (Eq. 4d, Ch. IX)

$\iota^s = 81;0;0;1$  = quantic number (cf. S. 42, 43, Ch. XI)

This formula describes a complicated situation such as a death rate,  $\%P$ , per month,  $(\%PT^{-1})$ , in each of a number of plurals,  $p$ . These rates are distributed by age periods,  $T^{-1}$ , the number of periods being  $\iota$ . The age periods do not have equal intervals between their terminal dates, i.e., they have more than one difference between initial dates,  $\iota'$ , and terminal dates,  $\iota$ , among the periods. This plurality of dates is denoted in the script by the unprimed small letter denoting an aggregation. Finally these mortality-by-age distributions are repeated for a series of years,  $\iota$  in number, subdivided into months,  $\iota'$ , between a single specified date,  $\iota'$ , and a single specified terminal date,  $\iota$ .

The quantic digit for an acceleration is 8. The complement from ten, of the exponent  $-2$  is used (as noted above in Chapter II when presenting the standardized notation) in order to avoid minus signs in the quantic number. This usage is comparable to the convention for writing the characteristic of a logarithm. For the combination of an acceleration and a duration, the digits 8 and 1 are used in the quantic number to denote the combination of what the digits 8 and 1 denote.

The above descriptive formulae are useful for the usual societal data where change may go on in jumps and not necessarily in an even flow. Wherever the change is an even flow describable by the equation of some curve, the differen-

tial coefficients of Calculus provide a far more refined measure of velocity than the usual averages of statistical series :

$$(I)_v = \frac{dI}{dT} = \text{a velocity, the differential coefficient of an indicant with respect to time} \quad (\text{Eq. 5, Ch. IX})$$

This states that velocity at any moment is the differential coefficient of the change with respect to time. Acceleration is given as :

$$(I)_A = \frac{d(I)_v}{dT} = \frac{d^2I}{dT^2} \quad (\text{Eq. 6, Ch. IX})$$

which is the second differential coefficient of the change with respect to time. (For a simple explanation of Calculus for unmathematical laymen, see Ref. 73.)

4. Another class of temporal phenomena is that denoted by a quantic digit of +2. This class comprises correlations between two series of durations, as when ages of husbands are cross-classified with ages of wives and with frequencies of such couples, or percentages of divorce, or fertility, or other indices, in the cells. See S. 9, Ch. XII for predictions of fertility from such a correlation matrix. See also S. 4, Ch. II, the length of unemployment cross-classified with ages, with consequent probabilities. Situations with a temporal exponent of two, whether plus or minus two, involve "secondary" temporal indices. "Secondary" means a multiplicative repetition of the index, either in a scalar product, or in an ordinary arithmetic product in a numerator or denominator :

$$T^{+2} = T^{+2}, T^{-2} = \text{a secondary temporal index} \quad (\text{Eq. 7, Ch. IX})$$

5. This formula in generalized form for any descript is :

$$I_s^2 T_s = T_s \quad (\text{Eq. 8b, Ch. IX})$$

6. Our conception of time is the usual solar time marked by annual and daily cycles of the sun and earth, and arbitrarily subdivided into hours, etc., by clocks. All sorts of units of time are possible, from vaguely delimited ones such as "2 cigarette smokes" or "an aeon," or "infancy, childhood, adolescence, adulthood," to precise ones such as minutes and light years. All these are interconvertible and can be re-expressed in solar units of time (within limits of accuracy proportionate to the constancy of the unit used). S-theory does not deal with non-scientific concepts of time, such as Bergson's concept of an unfolding of possibilities with no notion of an extension capable of being expressed either in cardinal units as a multiple of some standardized extension, or in ordinal units as a series of events in sequence.

7. These variants of the period script are shown in the following formulae :

$$t : u : v T = \text{a hierarchy of time periods ; time divided into } t \text{ periods, each subdivided into } u \text{ periods, and each further subdivided into } v \text{ periods} \\ (\text{See S. 1, Ch. IX}) \quad (\text{Eq. 9a, Ch. IX})$$

$$t, u T = \text{a duration subdivided into 2 alternative sets of units, or 2 sets of simultaneous periods} \quad (\text{Eq. 9b, Ch. IX})$$

8. The standardized notation of the date script may be noted by the advanced student. (See Eqs. 1b, 1c, 2e, 3d, 4d, Ch. IX.) In addition to these, there are :

${}^i\text{T}^{+1}$  = an aggregation of periods not in sequence, i.e., each period having limiting dates corresponding to it. The periods may be overlapping, or may be in sequence but non-consecutive (Eq. 10, Ch. IX)

${}^0\text{T}^0$  = the present moment (Eq. 11a, Ch. IX)

${}^{+0}\text{T}^{+1}$  = the past, an initial date up to the present (Eq. 11b, Ch. IX)

${}^{-0}\text{T}^{+1}$  = the future, the present date on to a terminal date (Eq. 11c, Ch. IX)

${}^{+0}\text{T}^{+1} : \text{P}$  = a distribution of people by age, i.e., from birth dates to a common present (Eq. 11d, Ch. IX)

${}^{+i}\text{T}^{+1} : \text{P}$  = a distribution of people by longevity, i.e., from varying birth dates on to varying dates of death, the differences being grouped into equal class-intervals,  $i$  in number (Eq. 11e, Ch. IX)

9. The geometric interpretation of temporal dimensions is very simple. The class script denotes the number of temporal vectors—usually only one and seldom exceeding two—in one S-situation. The period script denotes line sects ; and the date script denotes points on a temporal vector. Negative exponents denote division by a duration, in order to alter the units of an index to units per period. Since division of a vector by a vector is ambiguous, division of vectors is ruled out in mathematics. (Tensor theory, however, develops a form of vectorial division with special rules.) Quantities which are divisors must, therefore, be considered as scalar quantities and not as vectors. Hence T, with a negative exponent, is not a vector but a scalar, it merely alters the scalar units of the dividend vector. In sum, the general vectorial formula (Eq. 51, Ch. II) for a sector in S-theory holds as well for the temporal sector, as for any other.

10. Since more boys are conceived than girls, the ratio starts as larger than 1, but due to greater male mortality, drops to about 105 boys to 100 girls at birth and continues dropping to equality of the sexes usually by adulthood, and to an excess of females in old age. The rough linear relation above assumes a life span (T) of 70 years, a birth ratio of 1.05, and a final ratio of .95. Variations from this line are asserted in the dot which denotes a tendency to covary, a correlation, i.e., the sex ratio at any age subtracted from the sex ratio at birth and multiplied by 700, tends to equal that age.  $700(1.05 - \text{ratio at age T}) = {}^1\text{T}$ .

11. The principle as stated in Eq. 14 is a tautology and not an experimental finding or a metaphysical assumption. By defining a standard error as a standard deviation of the differentials of an index, the standard error necessarily varies as the square root of P, since it contains the root of P in its formula :

$$\sigma(\text{I}) = \sqrt{\Sigma \partial_i^2 / \text{P}} = \sqrt{\Sigma \partial_i^2} (\text{P}^{-.5}) \quad (\text{Eq. 14b, Ch. IX})$$

In deriving any formula for the standard error of any index, the procedure is to take the differential of the index by the rules of Calculus and consider this tiny deviation as a random fluctuation. These differentials are squared, summed, and divided by n (the number of samples), and the square root is taken to find

the standard deviation of these random fluctuations, or tiny errors, in a very large number of samples.

The experimental verification comes in at this point in the finding that the standard error thus derived by formula agrees closely with the standard deviation derived by computing the values of the index, if that index is actually re-determined from a large number of random samples of the same size. Barring observational errors (which are to be distinguished from sampling error) the residual, small, numerous errors, whose aggregate is called chance, are empirically found to be Gaussian, i.e., normal in their distribution, so that the standard error formulae furnish a satisfactory measure of the precision of the observed index, as far as sampling error is concerned. For a lucid and brief exposition of the theory of sampling error see Ref. 64, Ch. XVII.

**12.** For some highly developed examples of such indicators of culture complexes see the schedules for measuring rural hygiene, socio-economic status, public health departmental efficiency, the excellence of cities, in such references as 2, 12, 40, 76.

At the other extreme, for a simple situation of durations of qualitative attributes, consider such a quotation from the current press as: "Go to school to the age of eighteen, work from eighteen to sixty-five, retire and rest after sixty-five—that is the formula for solving the problem of unemployment as President Roosevelt enunciated it to the Young Democrats at Baltimore . . .," (Ref. 8) . . .  $I_i^0 : T^{+1}$  where  $|_i = 3$  school, work, and retirement

**13.** For a concise résumé see Appendices I and II. These give:

1. A glossary of terms and symbols.

2. Rules for writing S-formulae. With these the analyst is told what to do or look for, first, second, third, etc., until the recorded situation is symbolized in a descriptive formula and classified by its quantic number. This is an attempt towards providing operational definitions of S-theory concepts.

**14.** Some other common patterns of Brief-S formulae are:

"Full" "Brief"

${}^{\mathfrak{P}}P : (I) = {}^{\mathfrak{P}}(I) =$  a list of indices; for each person,  ${}^{\mathfrak{P}}|$ , a value of the index is stated (Eq. 21, Ch. IX)

${}_tT^{+1} : P = {}_tP =$  a population pyramid, i.e., for each class-interval of time,  ${}_t|$ , there is a corresponding population (Eq. 22, Ch. IX)

${}_iI :: {}_jI : P = {}_i :: {}_jP =$  a correlation scattergram (Eq. 23, Ch. IX)

$IT^{-1} = I_{/t} =$  a velocity; an indicator divided by time (Eq. 24, Ch. IX)

${}_{\mathfrak{T}}{}^{\mathfrak{T}}T^{-1} : (IT^{-1}) = I_{/tt} =$  acceleration; an index divided by time squared, in a series of periods (Eq. 25, Ch. IX)

${}_tT^{+1} : {}_{\mathfrak{P}}PT^{-1} = {}_{t/0}{}_{\mathfrak{P}}P =$  death rates by ages (Eq. 26, Ch. IX)

${}^{\mathfrak{P}}P :: {}^{\mathfrak{P}}P : I = {}^{\mathfrak{P}} :: {}^{\mathfrak{P}}I =$  the interrelations of P persons with each other (Eq. 27, Ch. IX)

$(PL^{-2}) = P_{/11} =$  densities; persons per unit of area (Eq. 28, Ch. IX)

15. In the Brief-S formulae note the product of a quality and a quantity. Thus in Eq. 19a, the  $|_p$  being singular is the quality "male" with no quantity asserted.  $L^3$  is the pure number of cubic units which, multiplied by the implicit attribute denoted by the class script  $|$ , "cranial," and divided by a second implicit product  $I^0, L^3 = L^3, =$  cc. of "body," becomes a ratio. The product of this ratio,  $L^3, , ,$  and  $|_p$ , makes it the ratio for males, or the male ratio. An adjective modifying a noun is the usual verbal form of expressing the product of a quality and a quantity, or of an attribute and an index ( $I^0(I) = (I, ,)$ ). (When the index is a dynamic one expressing action as in  $(IT^{-1})$ , it is expressed in words as a verb, and its modifying quality, the attribute, may be an adverb.)

Note that the plural descript denotes an aggregation in cross-scripting formulae also, so that in Eq. 17a there is a series of values of  $L_p^3$ , there being  $|$  items in the series, which are asserted to have a tendency to equal, item for corresponding item, the second series,  $|L_p^3, ,$  each augmented by 60 cc. Alternatively, this paralleling of two series could be expressed as a high correlation coefficient between them,  $(L_p^3) \cdot (L_p^3, ,) > .5$ . For series of only a few items, however, a correlation coefficient is not legitimately calculable.

## Chapter X

### CHANGE,<sup>1</sup> $T^{-1}$

#### I. CLASSIFICATION OF SOCIETAL CHANGE

All of the societal phenomena analyzed thus far in Chapters II through IX by means of the symbols of S-theory may undergo change in time. Accordingly, all the preceding chapters and their subtopics could be repeated in their dynamic aspects. Hitherto, the static aspects of these phenomena, represented by a point in the time dimensions,<sup>2</sup> have been sketched in confining the exposition to the  $T^0$  array of the quantic classification (see S. 33, Ch. II). The dynamic aspects, represented by the resultant vector compounded of a time vector and the vector of some other index, will be sketched in dealing with the  $T^{-1}$  array of that quantic solid.

#### A. Some Definitions

Before classifying the forms of societal change, a few distinctions may be useful between such overlapping concepts as, "change," "process," "growth," "evolution," "sequence," "tempo," and "velocity." "Change" will be used as the most general term covering all forms of something-varying-in-time. A "process" will denote a distinctive form of change symbolized by some particular formula in S-notation. A process usually connotes the more regular, continuous, and continuing change. "Progress" will denote a change desired by the population changing. It combines change with a value judgment as to its social desirability. By defining "progress" as change desired by the persons changing, it becomes objectively measurable by scientific technics.<sup>3</sup> Progress is thus relative to a group and date, and there are diversities of progress since there are conflicting desires between groups. "Progress" is the increasing of any positive value, as described in the tension theory in Chapter V. "Growth" is a change in a positive direction only, i.e., towards augmenting or increasing. Growth also connotes a life cycle of

the entity that is changing—a birth or beginning, a growth, a decline and death or cessation, in sequence. “Evolution” is defined more exactly below under the heading of compound processes. “Velocity” or “tempo” is the time rate of change, the ratio of the amount of change to the duration of the time in which it occurs,  $(I)T^{-1}$ . “Sequence” is the temporal order of events without necessarily having any continuity. In a change there is an intrinsic continuity, in that the entity changing is identifiable as that entity both before and after the change, whereas a sequence on three successive days of an earthquake, a birth, and a wedding, is not one but three changes, for there is no common entity running through this sequence.<sup>4</sup> \* For simplicity, the cross-scripted or Brief-S formulae will be used from now on. Each process will be defined by its symbolic formula and named, as far as possible, by some word which is current among sociologists at present and which most closely approximates the meaning of the algebraic symbols. Rigorous reasoning should be based on manipulating the symbols and the data they represent, and not on the “names” of these processes, since these names do not accurately define the processes. The names are used to help bridge the gap from familiar concepts to new ones. The new symbolized concepts are believed to be more objective, measurable, and therefore useful to science, but it is recognized that they are cryptic and at first forbidding to most readers. Hence the need of the more familiar terms as verbal crutches at first.

In choosing the names of these processes, a compromise between systematic regularity and current usage was adopted. Systematic regularity would call for a root word for each process with a prefix, such as “ad-” and “de-,” or “hyper-” and “hypo-,” to denote the direction of the process. In order to avoid neologisms, this regularity has often been tempered to use more familiar synonyms. Thus for the terms “increasing” and “decreasing” regularity would call for “creasing” = “accreasing” or “decreasing”—which seem too unfamiliar at present.

A further point on terminology is that the present active participle ending in “-ing” will be used to name processes, instead of the current nouns usually ending in “-tion.” Thus the terms “competing” and “associating” will be the names of the processes

\* For Eq. 1, Ch. X, see notes at end of chapter.

currently termed "competition" and "association." This terminology achieves two desiderata. It clearly denotes a process, an activity in time. It tends to avoid the "reifying fallacy," by which when something is named people tend to make a "thing," a noun, of it and forget that it is an acting, a verb.

Our process-formulae and their names are looked upon as descriptions of forms of acting. They are not explanatory principles, except insofar as it may be found that a correlation and a time sequence exist between a process and some other phenomenon which justify our saying that, with a specified probability, the phenomenon is a consequence of the process.

In the second place, the participle permits distinguishing the process from the relationship of the parties involved in the process. The relation may often be static. Thus parties who are competing may be in an unchanging relationship of competition to each other. The noun form will be used hereafter to denote the relationship, the participial form to denote the societal action or process. (Cf. Eubank's two categories of "change" subdivided into "action" and "relationship," Ref. 25, Ch. XIV.)

### B. Bases of Classification

The phenomena of societal change may be systematized by using five bases of classification successively. The first basis is the degree of *compounding* of symbols by which changes are classified into the single and the aggregated. A single change is one defined by a single index in the formula. A compound change is defined by an aggregation of indices—double, triple, etc., according to the number of indices in its formula.

$\text{t(I)}$ , = a single process (Eq. 2a, Ch. X)

$\text{t(I)}_i$  = an aggregated process (Eq. 2b, Ch. X)

Single processes will be expounded first in sections II and III of this chapter, and then aggregated processes in section IV of this chapter.

A second basis of classification is the *sector*. Under single changes:

populational processes =  $\text{tP}^p$  = "peopling" (Eq. 3a, Ch. X)

indicatory processes =  $\text{tI}^i$  = "indexing" (Eq. 3b, Ch. X)

spatial processes =  ${}_tL^1$  = "spacing" (Eq. 3c, Ch. X)

durational processes =  ${}_tT^{+1}$  = "timing" (Eq. 3d, Ch. X)

will be discussed in turn.

A third basis of classification is a corollary of the preceding basis. It is the number of sectors that are combined. A binary term is a product of one index and T<sup>-1</sup>. A ternary process is a product of change in two sectors (as IP<sup>-1</sup>T<sup>-1</sup>, a per capita change, or T<sup>+1</sup>PT<sup>-1</sup>, a change of man-hours). A quaternary process is a product of change in three sectors (as in PT<sup>+1</sup>L<sup>-1</sup>T<sup>-1</sup>, man-days of labor per mile of roadway built per year). A quinary process combines change in all four sectors. Binary processes will be called "simple," ternary and higher processes will be called "complex." Both are "single" in contrast to the aggregated processes. Section II of this chapter deals with the simple single processes, and section III with the complex single processes.

A fourth basis of classification is the *exponent*, the degree of operational thoroughness with which the situation has been observed and developed. On this basis the index changing may be:

nullary—with zero-degree exponents—

qualities changing  ${}_t(I^0)$  (Eq. 4a, Ch. X)

primary—with first-degree exponents—

quantities changing  ${}_t(I^1)$  (Eq. 4b, Ch. X)

secondary—with second-degree exponents—

relations changing  ${}_t(I^2)$  (Eq. 4c, Ch. X)

tertiary—with third-degree exponents—

further relations changing  ${}_t(I^3)$  (Eq. 4d, Ch. X)

Within each sector the processes defined by the successive exponents will be expounded in turn.

The fifth basis of classification is the *direction* of change within a dimension. Every process may go towards increasing or decreasing whatever is changing. For uniformity, increasing will be denoted always as positive, and decreasing as negative, taking the initial date of observation as the zero point of reference. The Brief-S formulae are:

${}_t(I)$  = changes in any index in the series  
of periods,  ${}_tT$  (Eq. 5a, Ch. X)

$+{}_t(I)$  = increasing of any index in the  
series of periods,  ${}_tT$  (Eq. 5b, Ch. X)

$-\iota(I)$  = decreasing of any index in the series of periods,  $\iota T$  (Eq. 5c, Ch. X)

$(I)_{,\iota}$  = velocities of change in the series of periods,  $\iota T$  (Eq. 5d, Ch. X)

These five bases of classification formally explore the deductions from the basic matrix equation defining S-theory. They are offered as fitting a large part of current concepts and data in Sociology and in the social sciences more generally, and as serving to systematize them in an orderly framework. Other classifications<sup>5</sup> of societal change will, of course, be needed for special purposes. Further extension of the bases here offered may be made, as in subclassifying the vast field of indicatory processes—perhaps in some such way as the classification of indicators by content offered in Chapter III.

## II. SINGLE PROCESSES, $\iota(I)$ ,

The single processes are those definable by a single index. If it is possible to describe the dynamic phenomena by a single index, that set of phenomena is defined as a single process. If more than one index is required, it is defined here as a compound process. The single processes may be conveniently subclassified by sectors beginning with the populational sector.

### A. Populational Processes, $P^p$ = "Peopling"

The populational processes are those in which the number of parties changes in some way in time. They will be subclassified according to the exponent as nullary ( $\iota P^0$ ), primary ( $\iota P^1$ ), or secondary ( $\iota P^2$ ), processes.

#### 1. SOCIATING, $\iota P^0_{\Sigma p}$ = A NULLARY POPULATIONAL PROCESS

Whenever the plurels in a population are non-overlapping groups of one kind, increasing of their number is a measure of the process of "dissociating," and decreasing of their number measures "associating," as ordinarily understood by sociologists, for whenever the groups are definite, effective dissociation must increase their number. Merging of several corporations into one, federating of churches or of states, the membership of several unions being collected into one union alone, are examples of associating as defined here. The separation of Ulster and the

former Irish Free State, of the Northern and the Southern Baptists, of the old Standard Oil Company into the Standard Oil of New York, of New Jersey, etc., are examples of dissociating. These processes may be symbolized as:

$$-{}_t|_{\Sigma_P} = \text{"effective associating"} \quad (\text{Eq. 6a, Ch. X})$$

$$+{}_t|_{\Sigma_P} = \text{"dissociating"} \quad (\text{Eq. 6b, Ch. X})$$

$$\pm{}_t|_{\Sigma_P} = \text{"sociating"} = \pm{}_t(P^0)\Sigma_P \quad (\text{Eq. 6c, Ch. X})$$

(See S. 2 and 10, Ch. X)

where  ${}_t|_{\Sigma_P}$  denotes the change in the number of groups (i.e., interacting plurels) with non-overlapping membership during the time T. These formulae measure the end result, the effective sociating, which culminates all the complex behavior tending towards sociating. The sociating situation can also be symbolized as a changing list or aggregation of plurels,  ${}_t|_P$ . This aggregative change is *measured* by the operation of *counting* the list, as symbolized by their sum,  ${}_t|_{\Sigma_P}$ . The formulae for this process (as for most others in this chapter) were inductively derived by an analysis of the essential properties of "association-dissociation," as currently described in the sociological literature. The essential property, "change in the number of groups," was then written in a formula by the usual rules of S-notation. It could also have been deductively derived.<sup>6\*</sup>

Any process may have its formula deduced from the basic S-theory equation by manipulating the scripts so as to consider any particular combination of scripts. New processes, not hitherto recorded in the sociological literature, may readily be derived and their properties stated as given by any unusual combination of scripts. *Every equation and S-situation in this volume can be viewed as a deduction from the S-theory equation (Eq. 50, Ch. II), i.e., as a particular permutation and combination of its scripts and operators. In this sense the forms of all societal phenomena as expressed in S-symbols are deducible from a single equation.*

## 2. POPULATING, ${}_tP^1$ = PRIMARY POPULATIONAL PROCESSES

Change in the numerical size of any population studied is simply defined by:

$$\pm{}_tP = \text{Populating} \quad (\text{Eq. 9a, Ch. X})$$

\* For Eqs. 7a-8c, Ch. X, see notes at end of chapter.

$+_tP$  = Adpopulating (Eq. 9b, Ch. X)

$-_tP$  = Depopulating (Eq. 9c, Ch. X) <sup>7</sup>

(See S. 10, 11, 12, 13, 34, 50, 51, and 55, Ch. X)

This is simply the lengthening or shortening of the population dimension. "Adpopulating" and "depopulating" may serve as general terms where, in particular connotations for particular types of plurels, such terms are commonly used as:

conscripting	and discharging	in a compulsory Army population
enlisting	and deserting	in a voluntary Army population
hiring	and firing	in a business population
joining	and leaving	in a fraternal population
membership gains	and membership losses	in a membership population
births	and deaths	in a living population
immigration	and emigration	in a regional population
registering	and striking out	in a recorded population
entering	and going out	in an attendance population
including	and eliminating	in a sampled population
enrolling	and resigning	in a voluntary registered population
conquering	and losing	in a colonial population
matriculating	and graduating	in a school population
marrying	and divorcing	in a married population
adopting	and disinheriting	in a family population of two generations
attracting clientele	and losing clientele	in a patronage population
baptizing	and excommunicating	in a church population
naturalizing	and denationalizing	in a national population
affirming allegiance	and denying allegiance	in a religious or political population
embarking	and disembarking	in a ship population
signing in	and signing out	in a hotel population

and a host of other such terms.

$\pm_tP$ , with the legend specifying what P is in the particular situation, states any and all of these processes. This illustrates a function of a systematic Sociology, namely, to describe phenomena in terms which are general to all the social disciplines, each of which may have its own forms of these phenomena and its own terms naming them.

As with all primary processes, the presence of class scripts does not alter their nature. The fact that the population studied may

be divided up into many pleurels,  $P_p$ , or among many characteristics,  $P_i$ , makes no difference as long as the size of the total population is moving in one direction, upwards or downwards.

Next, consider the pair of subprocesses defined by the tension theory of societal action outlined in Chapter V. One subprocess is delimited by the changing size of the population which desires a particular value,  $P_v$ . When its intensity of desire grows beyond the bounds of competition (for which see below), it breaks over into conflict.

In conflict the opposed parties of the population,  $P$ , who desire a given desideratum,  $V$ , carry their striving for it to the point of trying to eliminate the opponent. The primitive case of conflict is where the caveman, finding his competitor for game so successful as to result in starving him, turns upon his competitor and tries to kill him.

When competition for national aggrandizement becomes too acute, it breaks over into war, by attempting to depopulate the opposing armies or nations. Competition in sport turns into conflict when, instead of struggling for exclusive possession of the desideratum desired by both parties, namely "victory," the players start slugging and trying to cripple each other to eliminate opponents from activity in the realm of that desideratum. As means to the end of getting more of the desideratum, each party develops a supplementary desire to decrease the population of the opponent. If this is the distinguishing characteristic of conflict, conflict may be defined by:

$$-_tP_v = \text{conflicting} \quad (\text{Eq. 10, Ch. X})$$

(See S. 55, Ch. X)

$-_tP_v$  is the decrease of the population that is struggling for the desiderata,  $V_v$ . The population has decreased, regardless of which party suffered most in the mutual attempt to eliminate the other from that particular arena of conflict, from the field of one set of desiderata. For a rigorous definition to cover this means that whenever conflict exists, Eq. 10, Ch. X tends to be true, and conversely, whenever Eq. 10, Ch. X is true, conflict exists. "Tends to be" is inserted, as obviously actual slaughter may take time, as when a country mobilizes, or a Kentucky feudsmen stalks his enemy for months. Also, obviously people may die

from disease and accidents, making Eq. 10, Ch. X true, but here there is conflict with germs or physical forces as the opponents. "Eliminating" might be a convenient general term for either or both of human and non-human opponents. This theory, however, is only concerned with death, or elimination from effective opposition in the field of some desideratum, which is *caused by other people*, for only this is *societal* conflict.

The opposite of conflicting might be termed "recruiting,"  $+{}_tP_v$ , which is increasing the population defined by some desideratum, or set of desiderata.

In genetic and regional plurels, such as nations, the most significant analysis of populating into its constituent plurels is to state it as the algebraic sum of births, deaths, immigrants, and emigrants.

$${}_tP = {}_tP_b - {}_tP_d + {}_tP_i - {}_tP_e \quad (\text{Eq. 11, Ch. X})$$

This equation is the basis of vital statistics. When corrected for the age and marital status composition of a population, it becomes the basis for prediction of population trends. (See S. 77, Ch. X; S. 1, Ch. XI; and S. 14, 19, Ch. XII.)

Accordingly, as a nation is underpopulated or overpopulated in relation to its natural resources and its technology, it can adopt policies tending to control its populating by accelerating or decelerating change of each constituent plurel. Thus birth rates can be decreased by disseminating contraceptive information, economic deterrents for large families, failing to provide maternity health facilities, etc., while increased birth rates can be achieved by national propaganda and pressures, which have been so effective in Nazi Germany. Death rates can be decreased by public health measures and education in hygiene and other ways, or increased by neglecting these, as well as by policies leading to warfare. Migration can be stopped by passport control, and other legislation and policing; it can be increased by economic incentives of subsidized travel, inducements of free land, low taxes, etc., as well as by persuasive propaganda attracting immigrants, or coercive persecution expelling refugees.

In fact, each of the constituents in Eq. 11, Ch. X could be analyzed into its subconstituents and so on towards reducing them all to measurable entities, which are largely under the con-

trol of human organizations. Thus deaths can be analyzed as the sums of deaths from each disease or cause in the International Classification of causes of death. Many of these can in turn be expressed as resultants of conditions subject to societal modification, such as the social formula for reducing deaths from typhoid, which is chiefly to isolate cases of typhoid, to inoculate everyone, and to purify sources of water and milk. Proportionately as these steps are taken vigorously, typhoid deaths tend to disappear. Similarly other causes of deaths can be and are being progressively reduced, as the medical sciences discover their etiology, and society organizes to act on such knowledge. (See S. 62, Ch. X; S. 7 and 19, Ch. XI.)

These remarks on the populating process may serve to suggest its modifiability in response to human desires. A more rigorous discussion of the valid principles in vital statistics, Eq. 11, Ch. X, may be found in Lundberg's *Foundations of Sociology*, Chapter XI, where the causes and correlates of human reproduction are reviewed.

### 3. INTERACTING, $tP^2 = \text{SECONDARY POPULATIONAL PROCESSES}$

#### a. General.

The third type of populational processes is definable by an exponent of 2. This includes the dynamic aspect of all interrelations studied in Chapter VII. Since those interrelations were limited by definition to relations of stimulus and response, active or suspended, their dynamic aspect is "interacting."

$$t(P :: {}^pI_p :: p) = \text{interacting} \quad (\text{Eq. 12a, Ch. X})^{8*}$$

(See S. 50, 51, 52, 56, and 57, Ch. X)

The interacting processes might be classified, just as interrelations were, on the basis of content (sectors) or form (of the matrix surface).<sup>9</sup> This is, perhaps, the heart of the field of Sociology, but it has been so little developed as yet that no full discussion will be attempted here. Technics for measuring the interacting of parties and a detailed classification of interacting should prove a major field of sociological research. As partial evidence of the fruitfulness of this field for research, two types of interacting will be explored in more detail. They are interacting where the cell

\* For Eqs. 12b-c, Ch. X, see notes at end of chapter.

entries are people (P) and money ( $I_M$ ), in processes currently called "mobility" and "the economic process," respectively.

*b. Mobility.*

The special case of populational interacting of groups which usually goes under the rubric of "mobility" will be elaborated from its static aspect as given in Eq. 7, Ch. VII. This involves the transfer of personnel between all pairs of non-overlapping groups in a population. The complete matrix is:

	$P_1$	$P_2$	$P_3$		$P_Z$	
$P_1$	$+P_{11}T^{-1}$ $-P_{11}T^{-1}$	$P_{12}T^{-1}$	$P_{13}T^{-1}$		$P_{1Z}T^{-1}$	
$P_2$	$P_{21}T^{-1}$	$+P_{22}T^{-1}$ $-P_{22}T^{-1}$	$P_{23}T^{-1}$		$P_{2Z}T^{-1}$	$= (P_p :: P_p)T^{-1}$ = the mobility matrix
$P_3$	$P_{31}T^{-1}$	$P_{32}T^{-1}$	$+P_{33}T^{-1}$ $-P_{33}T^{-1}$		$P_{3Z}T^{-1}$	$= {}_tP_p :: P$ in Brief-S (Eq. 13, Ch. X)
$P_Z$	$P_{Z1}T^{-1}$	$P_{Z2}T^{-1}$	$P_{Z3}T^{-1}$		$+P_{ZZ}T^{-1}$ $-P_{ZZ}T^{-1}$	

where the small letter descript,  $|_p$  (rows), denotes the plurel losing personnel, and the capital letter descript,  $|_P$  (columns) denotes the plurel gaining personnel. The cell entries other than in the main diagonal denote the number of persons transferred. The cell entries in the main diagonal may be either:

- the number of persons leaving a plurel and returning to it again,  $P_{,,}$
- a summary of the row and column totals for that plurel in the form of either the algebraic sum to measure net mobility, or the absolute sum to measure gross mobility.

Net mobility measures the change in total number of members of a plurel, as in hiring and firing in a business, matriculating and graduating from a school, immigrating and emigrating from a region. Gross mobility measures the turnover, the total amount of adpopulating plus depopulating. Net mobility is the differ-

ence between the two populating processes, while gross mobility is their sum.

To summarize the amount of mobility in an aggregation of plurels, the standard deviation of gains and losses can be built up into an index of mobility as follows:

$$(.5(\%t \cdot P^2)_{\Sigma p})^{.5} = Mb_N = \text{index of net mobility} \quad (\text{Eq. 14a, Ch. X})^{10}$$

(See S. 10, 11, 18, Ch. X)

$$(.5p^{-1}(\%t \cdot P^2)_{\Sigma p})^{.5} = Mb_G = \text{index of gross mobility} \quad (\text{Eq. 14b, Ch. X})^{10*}$$

When the gains and losses are expressed as percentages of the initial population, their standard deviation in the p plurels (when expressed as a percentage of the standard deviation of the maximum possible mobility) is this index,  $Mb_N$ . Net mobility is zero when all plurels end up with the size of their populations unchanged; and net mobility is 100% when an initial monopoly by one plurel of all the population is reversed and ends up as a terminal monopoly of the membership by another plurel. This reversal of monopoly is the maximum possible mobility.

Gross mobility is zero only when no persons have changed plurels, rises to 100% when the whole population changes plurels once on an average, and may rise above 100% if the whole population on an average changes plurels more than once in the period studied.

Note that any gain or loss in the entire population studied is the populating process, and must first be subtracted, or eliminated by using percentages in place of absolute numbers, in determining the mobility which, by definition, is solely the *redistribution* of persons between the plurels within the entire population that is selected for study. Populating is measured by a mean, the statistical first moment; mobility is measured by a sigma, derived from the statistical second moment. Hence mobility is termed a secondary process, while populating is a primary one. Populating shows the movement of the whole population; mobility shows movements within the whole population studied.<sup>11</sup> †

Special terms to differentiate mobility among groups of diverse

\* For Eqs. 15a-b, 14c, and 16a-b, Ch. X, see notes at end of chapter.

† For Eqs. 17-21, Ch. X, see notes at end of chapter.

sorts are useful. Thus immigrating and emigrating refer to regional or geographic plurels.<sup>12</sup> \*

"Touring," "commuting," "visiting," "attending," etc., connote more temporary migrating. "Turnover" usually connotes employment groups. Then there are many vaguer terms in use, such as "shifting clientele," "fluctuating membership," "variable attendance," and the like. (For some mobility situations see S. 8 and 10, Ch. X.) These terms can be given precise operational definitions whenever needed on the pattern of the index of mobility.

c. *The economic process,  $\frac{1}{t}M_{p::p}$ .*

As another sample of the vast field of interactions awaiting exploration by more exact and powerful matrix technics, consider the economic process, which is an aggregation of all economic acts. This process can be systematically set forth in the usual "interacting matrix" as follows: the matrix will be a fourth-degree one, that is, it has four plural descripts ( $|_{\Sigma S} = 4$ ).

Let every buyer be listed as a column heading and every seller as a row heading. Completeness requires as many arrays (rows or columns) as there are parties in the population studied. Every cell contains an account during a specified period between the row and column pair of parties. The indicator in the cell is a monetary unit,  $M$  (more convenient than  $I_M$ ).

Usually the monetary exchanges (i.e., interactions of the parties) are considered during a period, and hence are velocities of the economic process. The matrix for one period and with a single indicator of exchanges, namely money, can be represented in a second-degree square matrix, thus:

		Buyers, $P_p$					$\updownarrow$ Direction of flow of money
		$P_1$	$P_2$	$P_3$		$P_Z$	
Sellers, $P_p$	$P_1$	$\pm \frac{1}{t}M_{11}$	$\frac{1}{t}M_{12}$	$\frac{1}{t}M_{13}$		$\frac{1}{t}M_{1Z}$	$= \frac{1}{t}M_p : p$ in Brief-S $= P_p :: P_p : (I_M T^{-1})$ (Eq. 23a, Ch. X) $ ^s = 9;1;0;2$ $ _{\Sigma S} = 2$ the economic process, i.e., exchanging or in- teracting measured in monetary units
	$P_2$	$\frac{1}{t}M_{21}$	$\pm \frac{1}{t}M_{22}$	$\frac{1}{t}M_{23}$		$\frac{1}{t}M_{2Z}$	
	$P_3$	$\frac{1}{t}M_{31}$	$\frac{1}{t}M_{32}$	$\pm \frac{1}{t}M_{33}$	.	$\frac{1}{t}M_{3Z}$	
	$P_Z$	$\frac{1}{t}M_{Z1}$	$\frac{1}{t}M_{Z2}$	$\frac{1}{t}M_{Z3}$		$\pm \frac{1}{t}M_{ZZ}$	
		$\longleftrightarrow$ Direction of flow of commodities					

\* For Eqs. 22a-b, Ch. X, see notes at end of chapter.

If a series of periods is considered, the period subscript ( $\tau$ ) becomes a plural one ( $\tau$ ) expanding the matrix to a third degree, which is visualizable as pages of an account book, each page being the matrix equation above. If, further, the record is to be kept separately for every commodity exchanged, the monetary indicator assumes a plural class script and the matrix is expanded in the fourth degree. This can be visualized as an account book for each commodity.<sup>13</sup> \*

In the matrix, money (which here includes gold, paper, credit documents, and all forms of media of exchange) flows vertically, up and down from row to row, from seller to seller, since by definition a seller is a party who receives money and gives commodities. Commodities, in the matrix representation, flow horizontally, left and right, from column to column, from buyer to buyer.<sup>14</sup> When all forms of exchangeable commodities are considered, the row totals must balance the column totals, provided proper entries are made in the main diagonal cells.

The monetary entries in the main diagonal cells are the heart of the economic process. Let a positive entry in the diagonal cell denote the total income to that party and a negative entry denote the total expenditures, while the algebraic sum is that party's profit or loss in the period specified.<sup>15</sup>

The effort in a capitalistic economy is to have positive entries (in money terms) exceed negative entries, in order to make a profit. The positive entry in a diagonal cell is the sum of the row entries; the negative entry is the sum of the column entries. Thus, the balance in any diagonal cell equals the sum of that party's row and column, which is the fundamental accounting equation of any business:

$$(+_{\tau}M) + (-_{\tau}M) = \pm_{\tau}M = \text{accounting equation of business} \quad (\text{Eq. 24, Ch. X})$$

income — expenditure = profit or loss in each period of a series

The point of this exposition of the exchange matrix is this hypothesis: every recorded set of economic data in the universe is a selection from, or a function of, this exchange matrix—the economic process symbolized in Eq. 23, Ch. X. This may be

\* For Eq. 23b, Ch. X, see notes at end of chapter.

called the "exchange matrix hypothesis." It claims that every account kept by any business, every inventory, every financial statement, every tariff schedule, all economic statistics in short, are derived from this matrix equation. The author has not yet found any quantitatively recorded economic data that could not be so analyzed.

For a few cases in evidence consider the following:

1. A shopkeeper's journal of sales and purchases is the aggregation of items, all of which are in his one diagonal cell in Eq. 23a, Ch. X (since here he does not keep a separate account for each buyer from him or each seller to him). This cell is expanded for  $I_m$ , the  $|_m$  different commodities exchanged and their prices  $I$ , and thus the cell, a zero-degree matrix, becomes an array, a first-degree matrix. If further, the situation  $S$  is defined as the journal for more than one day, then the  $I_m$  array is further expanded in the time dimension,  $\text{tT}$ . Such a journal is a second-degree matrix of the itemized exchanges of one party for a series of periods.  
 $\text{'P} :: \text{P}_{\Sigma p} : \text{t}(\text{IT}^{-1})_m$ .

2. A national budget is but the statement of future expected revenues and expenditures of one national plurel for a period.  
 $\text{P}_{p'} :: \text{'P}_{\text{P}} : \text{'t}(\text{IT}^{-1})_m$  where  $|_{p'}$  = the Government.

3. A table of imports and exports, such as S. 37, Ch. XI, is the row and column for one party in Eq. 23a, Ch. X expanded for a series of years and for subtypes of trade such as re-exports and transit trade. The indicators need not be monetary ones, as tons, bushels, yards, bales, carloads, percentages, index numbers, and other units may be substituted for  $I_m$ .

4. A large department store, where every sale is recorded by punching on a card the purchaser's charge account, the identification of the salesgirl, the date, the commodity, and the price, may have its thousands of accounts balanced by means of a Hollerith tabulating machine within a few minutes after the afternoon's closing hour. The tabulator can be set to summate the patrons' accounts to date ( $\text{P}_{\text{P}} : \text{'I}_m$ ); or the salesgirl's sales for the day ( $\text{'P}_{\text{P}} : \text{IT}_m^{-1}$ ), as the employees are one type of party who sell their services to the store; or the stock on hand of each inventoried class of commodity ( $\text{'I}_m - \text{'I}_m$ ); or the total sales for the day ( $\text{IT}_m^{-1}$ ).

5. The economic subprocesses of economic competing, con-

flucting (bankrupting), co-operating, progressing, dissimilarizing (division of labor), associating (merging), and the rest, are all measured and defined by indices derived from the matrix Eq. 23, Ch. X. Thus competing is a sigma of the diagonal cells, measuring net gains and losses in relative standing of all the competing parties in a period. Thus conflicting is eliminating of competitors by bankruptcy and business failure, reducing the number of parties,  $|\Sigma_P$ . Co-operating is increasing the mean value produced,  $\Sigma(IT^{-1})_M/P$ . Dissimilarizing is increasing the kinds of commodities exchanged,  $+_t|\Sigma_m$ . Associating is merging, buying up, forming a holding company, or other forms of complete, or hierarchical, reduction in the number of competing parties without elimination of any,  $-_t|\Sigma_P$ .

In concluding this discussion of the economic process, it is submitted that the exchange matrix, Eq. 23, Ch. X, is a very inclusive and exact mathematical representation of all economic exchanges, and a basis for derived data. It is a basis for all summaries and special economic analyses. By systematic varying of its scripts new forms of economic analysis can be deduced. The measurement of economic competition, to be expounded in the next section, is just one such possible deduction.

In turn, this exchange matrix is but a special case where the indicators are monetary ones, or equivalents, of the more general interacting matrix, Eq. 12a, Ch. X. This, in turn, is but a particular combination of the indices and scripts of the basic S-theory equation. Thus, again the S-theory is seen as a highly general formulation of societal phenomena from which the whole of such large fields as economic data <sup>16</sup> can be deductively derived as special cases.

### *B. Indicatory Processes, $I^i$ = "Indexing"*

In this systematic exploration of processes defined by the four indices P, I, L, T, each in turn, the populational processes having been described, the indicatory processes will next be expounded. Subclassifying these again by their exponents yields nullary, primary, secondary, and tertiary processes, each of which will be exposed in turn.

## 1. DIFFERING AND SIMILARIZING, $\mathcal{I}_i^0 = \text{NULLARY INDICATORY PROCESSES}$

The simplest process here is the qualitative changing of a characteristic. We say: "sweetness" has changed to "irritability" in someone's personality; the fashion in hats has shifted from style A to style B; ice cream was substituted for pie as dessert; the Swedish-speaking immigrants became English speaking; his concept of God developed from an anthropomorphic one to a system of supreme values; and an infinity of other changes in qualities or attributes. These are all symbolized by a change in an attribute, or a hierarchy of attributes, and may be called "differing":

$$\mathcal{I}_i^0 = \text{"differing," a changing quality} \quad (\text{Eq. 25a, Ch. X})^{17} * \\ (\text{See S. 20, 21, 22, 30, 35, Ch. X})$$

Next in dealing with an aggregation of qualities, if the change is towards greater similarity, i.e., converting them into fewer distinguishable qualities, this process may be called "assimilizing." Its opposite is conventionally known as "differentiating," but for uniformity of terms (as well as to avoid the mathematical connotation of differentiating in Calculus), the term "dissimilarizing" is suggested.

$$\pm \mathcal{I}_i^1 = \text{"similarizing"} \quad (\text{Eq. 26, Ch. X})^{18} \\ = \text{dissimilarizing and assimilizing} \\ (\text{See S. 10, 19, and 23, Ch. X})$$

Thus in S. 23, Ch. X for each of the sixteen commodities listed, the Department of Commerce's proposed reduction in the number of styles manufactured represents an assimilizing process for each of these sixteen qualitative characteristics.

## 2. INDICATING, $\mathcal{I}_i^1 = \text{PRIMARY INDICATORY PROCESSES}$

Probably the largest class of processes in the S-theory classification is that defined by a changing indicant. The variety of sub-processes here is enormous. For one suggestion of the extent of the field, consider the classification of indicators by content in Chapter III. The majority of societal processes described by sociologists to date seem to be reducible to subtypes of change in

\* For Eqs. 25b-f, Ch. X, see notes at end of chapter.

an indicator (or an aggregate of indicators),<sup>19</sup> \* much of which is change in indicants.

The most general process of significance for Psychology and Sociology alike is "behaving," the behavior of a human being. Any observed unit of behaving, a behavior pattern, can be symbolized in S-notation as an organized series of activities of a person by means of the formula:

$$\underset{\text{(full formula)}}{\mathcal{P} : {}_t(I^{0,1,2}T^{-1})_{i,j,---:z}} = \underset{\text{(Brief-S formula)}}{\mathcal{P}'I_s^i} = \underset{\text{behavior pattern}}{\text{behaving, a person's}} \quad (\text{Eq. 29a, Ch. X})$$

This asserts that for a typical person,  $\mathcal{P}'$ , there is a temporal series,  $t$ , of aggregated and hierarchically organized,  $|_{i,j,---:z} (= |_s)$ , activities,  $IT^{-1}$ , of qualitative, of quantitative, or of correlated types,  $|^{0,1,2} (= |^i)$ . By deleting the prime and the underlining of the person script, the formula shifts from typical behavior to the behavior of each person in a *distribution* of individual differences (Eq. 29b, Ch. X).

By shifting from the person script to the plural script the formula describes the behaving of plurals, which is more completely in the field of Sociology. One form of the formula for a typical unit plural observed on a series of dates now becomes:

$$\mathcal{P}^t I_s^i = \text{a plural's behaving} \quad (\text{Eq. 29c, Ch. X})$$

To make the formula determinate, to substitute numbers and definite qualities for the generalized algebraic letters in a specified behavior pattern, is the usual function of scientific observing. As always, a formula, as the word denotes, is a brief expression of the *form* of phenomena, leaving the particular content which has that form as a further subvariable. The task of psychological research upon any specified behavior pattern is to list the objectively distinguishable items of action ( $|_i$ ), to discover their hierarchical organization into constituent patterns and more inclusive patterns ( $|_s$ ), to determine these items qualitatively, then quantitatively if possible, then to determine their intercorrelations as far as observable ( $|^i$ ), to repeat all this for the successive periods or phases of the total period of the pattern observed ( $t$ ), and finally to specify the conditions in defining the typical person in whom this behavior pattern is observable ( $\mathcal{P}'$ ), as in classifying him as "normal or insane," "infant or adult," with such and

\* For Eqs. 27-28, Ch. X, see notes at end of chapter.

such attitudes, or mental sets, i.e., implicit attributes,  $I^0$ , in the situation observed, etc. As always the formula is not a substitute for careful observing and devising instruments for observing. It merely expresses observations in formal, i.e., standardized or generalized, fashion, thus fulfilling the parsimony aim of science. It then enables comparisons, compounding, and further treatment, whether mathematical treatment in manipulating the symbols to discover new relationships, or applied logical treatment in deducing further specific behavior from a given formula.

The formula, Eq. 29, in common with most formulae in this volume, can be considered as a methodological hypothesis. The hypothesis might be worded as, "This symbolic formulation of . . . ('a behavior pattern,' in this instance) can be a useful tool for more exact observing, generalizing, and relating the phenomena which it symbolizes." This hypothesis will be disproved if, after fair trial, sociologists are unable to use it fruitfully; it will be proved in proportion to their accumulating fruitful applications of the formula to data.

Without attempting an exhaustive classification of indicating processes, a few which are of paramount importance to human beings because they deal with human desires and values (*desiderata*), will be outlined as suggestions for further refinement. These are the three processes defined by change in the three indices of intensity of desire, *desiderata*, and tension ( $D$ ,  $V$ , and  $E$ ) in the tension theory, the static aspect of which was presented in Chapter V. *a. Valuating,  $D$ .*

The tension theory of societal action postulated that people desire *desiderata*, and, therefore, behave so as to increase the positively desired *desiderata*, and also so as to decrease the negatively desired ones.

In Eq. 34, Ch. V,  $PD = VE$ , a population ( $P$ ), defined as those desiring a specified *desideratum* (i.e., object of desire) whose available amount is denoted by  $V$ , shows an average intensity of desire,  $D$ , for that value with a resulting societal tension,  $E$ , which is measured by the ratio  $PD/V$ . This is definitional; the hypothesis built on these definitions is that societal tension thus defined tends to cause (i.e., correlates with on allowing for a time lag) societal actions relative to that *desideratum*.

Consider the meaning of change in each of these defined quan-

titles: D, V, and E (which are symbols substituted for brevity for the indicants  $I_D$  = average intensity of desire,  $I_V$  = the available quantity of the desideratum, and the index  $(I)_E$ , the resulting tension). The intensity of desire for an object is a valuation of that object. Intensifying a desire may be termed "evaluating"; its opposite, "devaluating"; and the pair, "valuating."

$$\pm_t({}^{2p}I_D)/P = \pm_t D = \text{valuating, change of average intensity of} \\ \text{desire} \quad (\text{Eq. 30, Ch. X}) \\ (\text{See S. 10, 32, 42, Ch. X})$$

As will be developed more fully in Chapter XI on Forces, this process of intensifying human desires is the chief type of societal force, the cause of human behavior. Its subvarieties are a major part of the field of Psychology, as suggested in such concepts as "drives," "motivation," "urges," "emotions," "glands," and other topics underlying the dynamics of behavior. Desires in complexly conditioned structures are a large part of such concepts as sentiments, wishes, psychoanalytic complexes, ideals, and loyalties. As usual, the exact definition of this valuating process should not be expected in verbal terms of synonyms and description, but in operational terms of attitude tests and other scales with their manuals of instructions for administering them. Progressively, as such technics for measuring human desires in specific situations become developed, the valuating process and its subforms may be expected to become first determinate, and later, from accumulating knowledge of its correlations, increasingly predictable and controllable.

One subprocess of devaluating occurs when the desires of the parties are mutually incompatible. This is "accommodating." When a conquered people as in Abyssinia cease struggling against Italy and accept the inferior status, they are accommodating, i.e., devaluating the value of "independence," which has proved incompatible with the desire of the stronger conquerors. When their desires for independence increase again, evaluating that value is going on and tends toward an uprising. Accommodating may be one-sided devaluating, resulting in subordinate-superordinate statuses of the parties; or in mutual devaluating, as in compromising with resultant equal status.

$$-{}_tD_{V:P::P} = \text{"accommodating"} \quad (\text{Eq. 31, Ch. X})$$

This states that intensity of desire,  $D$ , is decreasing ( $-D$ ) in the series of periods,  $\{t\}$ , among the interacting parties,  $\{p::P\}$ , with respect to a specified set of values, or desiderata,  $\{v\}$ .

In Economics this accommodating is made by the mechanism of the price rising (according to the marginal utility theory). As the price rises, some of those desiring the goods stop bidding, i.e., cease to be part of the economic demand. These persons find that their desire (or "evaluation"), measured by what they are willing and able to pay, has fallen below the price, and so they are no longer "in the market." In social fields refugees entering a region, slum residents invading a superior residential district, foreigners and strangers in general, illustrate the accommodating process when incompatible desires are being reduced to a compatible point, which is that point at which the available desiderata are enough for all those desiring them. Each limited desideratum is devaluated in the sense of having a lower worth assigned to it; it is less cared for than formerly.

*b. Grading,  $\{V\}$ —Progressing and regressing.*

The enrichment of people with more of whatever are their desiderata is progress to them, and their impoverishment with loss of these desiderata, or increase of their negative desiderata, is regress to them. Such a definition of progress removes it from the realm of speculation or controversy and makes its determination a matter of scientific fact finding. Technics to determine what a people want will define progress to that people at that time. This is the positive point of view of science rather than the normative point of view of ethics. Progress may vary between plurels and periods, but with research the degree of universality of the characteristics composing it (the desiderata) can be progressively determined.

Increase of a people's desiderata may result from natural forces or from additional human effort. The former type may be called "accumulating," as in the formation of coal, forest growth from rainfall, and the spawning of fish in the sea. The other type where the intentional<sup>20</sup> interaction of people results in increasing a desideratum is the essence of "co-operating."

The opposite processes might be termed "decumulating," where the desideratum is decreasing without human effort; "mal-operating," where the desideratum is decreasing by means of

human effort; and "regressing," where either or both of decumulating or maloperating are going on. The process-pair, going in either direction, might be referred to by the roots "cumulating," "operating," and "grading." <sup>21</sup> \*

- $\pm_i V$  = grading (Eq. 34, Ch. X)  
           = change in the amount of a desideratum (object of desire)  
 $+_i V$  = progressing (Eq. 34a, Ch. X)  
 $-_i V$  = regressing (Eq. 34b, Ch. X)  
           (See S. 10, 19, 25, 26, 32, 37, 40, 41, 43, 54, Ch. X)

As the boundary between cumulating and operating is often hard to draw, a term such as "grading" for the combined process going in either direction is useful.

Co-operating, like other processes, is seldom found in society in a pure and isolated state. Usually several processes co-exist. A single process is simply an abstraction used by sociologists as an element for analysis and understanding of the situation and chosen by reason of its being a significant, but observable and measurable, characteristic in the total flux of societal development.

An example of the co-existence of simple processes is the business system of capitalistic countries (A.D. 1938). Along with the competitive attempt to draw business away from the other fellow, this fear of losing as well as desire to gain is a strong incentive to co-operation as here defined, i.e., to produce more  $V$  for all. Thus auto manufacturers are partly competing, partly dissimilarizing in offering diverse cars, partly accommodating in code agreements, and partly co-operating in producing more  $V$  (automobiles) for everyone to enjoy.

The co-operation may be of varying extensity (even in the realm of one desideratum that is desired with a given intensity), depending on the proportion of the desideratum,  $V$ , that is created by each person who is desiring and sharing it. Thus the co-operation is minimal when one person (such as an outside philanthropist) creates something only slightly desired by the people; it is maximal when every one of the persons desiring it contributes to its creation. This is the momentum of co-operating, i.e., the rate of achievement of a desideratum, multiplied by the number

\* For Eqs. 32-33, Ch. X, see notes at end of chapter.

of persons creating it, as defined later in this chapter in the section on complex single processes.

c. *Tensing*<sup>22</sup>,  $\pm E$ .

In the tension theory of societal action, equilibrium,  $E$ , was defined as being the ratio (in appropriate units) of the population times their average intensity of desire divided by the available quantity of the desideratum.  $E$  is thus the ratio of the total desire of a population for a desideratum to the amount of the desideratum available. In Economics it is the ratio of demand to supply (called the "economic value" of the commodity). As this ratio increases, the disequilibrium, or "tension," increases in that population in respect to that desideratum. Increasing tension may be termed "attensing." As the ratio decreases, a process which may be termed equilibrating, or "detensing," is going on.

$$\pm_t(PDV^{-1}) = \pm_t E = \text{tensing} \quad (\text{Eq. 35, Ch. X})^{23} *$$

(See S. 10, 27, and 42, Ch. X)

Thus the three conventional societal processes of "conflicting," "accommodating," and "co-operating" emerge as subtypes of three processes which are by definition all on the one continuum of  $E$ , the societal tension in respect to specified desiderata. Tension will increase either as the average intensity of desire for the desideratum increases ( $+_t D_v$ ), or as the per capita share of the desideratum decreases, because there are more sharers ( $+_t P_v$ ) or less of the desideratum ( $-_t V$ ). Tension will be eased ( $-_t E$ ), either as people desire the desideratum less intensely ( $-_t D_v$ ) in accommodating, or as the per capita share of the limited desideratum increases through the outcome of conflict reducing the number of sharers ( $-_t P_v$ ), or through co-operating in creating more of the desideratum to be shared ( $+_t V$ ).

Tensing is defined and measured by Eq. 35, Ch. X. Our "tension hypothesis" is that this tensing is a psychological cause, conscious or unconscious, of behavior. In plurals it is that which produces societal action. In asserting a causal connection between our tension,  $E$ , and societal action, the three necessary and sufficient conditions of causation must hold: (1) tensing must precede

\* For Eqs. 35a-b and 36a-b, see notes at end of chapter.

the action to ease it; (2) there must be a correlation (in a given population, of course) between tensing and acting; and (3) the attendant conditions must be specified in some duplicatable form.

An exhaustive examination of even the evidence available at present for testing this hypothesis will not be attempted here. It is a major research in itself. Only a few hints, suggesting the reasonableness of the hypothesis in order to interest others in exploring it, will be given in sketching roughly some current examples on an international scale.

Consider the Sino-Japanese situation up to 1938. A growing population in Japan was increasing tension there faster than increasing industrialization could increase economic desiderata. Thus the per capita purchasing power, or average standard of living, of the Japanese citizenry is alleged to have declined. Increase of P in Eq. 35 exceeded increase of V, resulting in increased E. Meanwhile the imperialist ambitions expanded. The ruling classes, including business leaders, built up the popular intensity of desire for expansion on the mainland of Asia. As the press and military-dominated government increased D, E increased still further, until a critical point was reached where conflict became the only way of easing the high tension. It was the only way remaining to the Japanese, when it is seen that neither technology nor trade organization could ease the tension by increasing the economic desiderata adequately, and that for reasons of prestige they were unwilling to devaluate (i.e., to decrease) their imperialist ambitions. Reasoning on the basis of Eq. 35 it may be predicted (in May, 1938) that the conflict will go on until terminated by one or more of the following factors:

- a. Large casualties (easing tension due to population pressure, plus setting up counter tensions towards peace, since the human lives destroyed are a desideratum to their families)

$$(-_tP_v) \bullet (-_vE_v) = r > 0 \quad (\text{Eq. 37a, Ch. X})$$

The equation states (in S-notation towards reducing the symbolized phenomena to measurement) that elimination of population tends towards detensing, i.e., they are positively correlated (r) after allowing for a time lag, as denoted by the periods  $_t|$  being antecedent to the periods  $_v|$ .

- b. Conquest of much of China (easing tension by increasing the desiderata,  $|_v$ , desired by the Japanese)

$$(+_tV_v) \bullet (-_vE_v) = r > 0 \quad (\text{Eq. 37b, Ch. X})$$

This equation states that detensing tends to follow co-operating, i.e., that detensing correlates positively with the prior increase which is achieved in what are desiderata to the Japanese.

- c. Waning of imperialist enthusiasm (easing tension by decreasing the intensity of desire for the desiderata at stake in the war)

$$(-_tD_v) \bullet (-_vE_v) = r > 0 \quad (\text{Eq. 37c, Ch. X})$$

This equation states that detensing (or equilibrating) tends to follow the devaluating of the desiderata which created the tension. This devaluating could conceivably come by choice of those in control to steer public opinion away from ambitions to expand, but it is more likely to come by the general growth of counter desires and tensions to stop casualties, heavy taxation, and war privations.

These three factors include in general all the possibilities as revealed formally by Eq. 35b, or less formally by a common sense analysis. To predict which factors will be operative, and in what proportions, requires measurement of these factors in intricate and complete systems, together with their recent trends for projecting into the near future, to an extent that has not yet been achieved by social scientists or statesmen. Their crude subjective estimation, however, is all that statesmen have at present for guiding national polity.

The function of the formal analysis in symbols is to facilitate the quantitative determination of the symbolized entities and their rigorous manipulation, and so increase society's ability to predict and control its future.

Consider next the Chinese situation.<sup>24</sup> The conflict was thrust upon them by Japanese demands and seizures which chiefly deprived China of some of its desiderata, i.e., decreased  $V$  ("mal-operating") and consequently increased tension in China ("attensing"). In 1931 China accommodated  $(-_tD_v)$  in the similar Manchukuo situation, but in 1937 not only was the loss of desiderata greater, but the intensity of desire for it ( $_tD$ ) and for resist-

ing the Japanese at all costs had increased. The tension in 1937 resulted in China's preferring conflict to further accommodation as the means of easing the tension.

For China the conflict may be expected to continue until:

- a. Lost territory is largely regained (easing tension by increasing the decreased desiderata)

$$(+{}_tV_v) \bullet (-{}_uE_v) = r > 0 \quad (\text{Eq. 37d, Ch. X})$$

This parallels Eq. 37b.

- or b. The desire to resist wanes (easing tension by accommodating to superior force)

$$(-{}_tD_v) \bullet (-{}_uE_v) = r > 0 \quad (\text{Eq. 37e, Ch. X})$$

This detensing through devaluating may be expected, if loss of life and of other desiderata and low probability of eventual victory set up counter tensions towards peace sufficient to produce devaluating of the war objectives.

Proportionately as Eq. 37c and e both take place a compromise peace will be the outcome, while if one goes on with less force than the other that party will be the victor in the war.

Other situations of international tension, whether eased by conflict, or treaties, or otherwise, may be similarly analyzed in terms of changes in the three independent factors, P, D, and V, and the resulting ratio, E. Industrial and interracial tensing, interreligious and domestic tensing, may be similarly analyzed towards increasing the precision of their observation and eventual control.

At the lower limit when the population reduces to one person ( $P = 1$ ), tensing becomes a process of individual psychology. Eq. 35 still holds whether for normal degrees of tension or abnormal tensions which are eased in paranoiac delusions and other psychoses. If a person keeps his desires within a bridgeable interval of the available amount of the desideratum, his tension will spur him to act and strive to achieve. If he indulges in extreme desires, as in daydreaming fantasies of himself as ruler of the universe, the resulting tension is eased by ignoring reality and developing a stimulus-proof delusion of grandeur. If he desires little and happens to receive an amount of a desideratum in excess of his desires, tension sinks to zero. An easy-going South-

ern negro, accustomed to a low standard of living, will, if he chances to receive a windfall of money, stop work and lie around until it is consumed, and until hunger and other appetites revive tension in him and drive him to work again.

A supply of the desideratum greatly in excess of the desire for it may even produce a negative tension, as the desire for more of the desideratum turns into aversion. Thus a well-fed person on being forced to eat more and more develops a temporary loathing for food.

In sum, for the individual the tensions causing his behavior may be analyzed into: (a) inner states, including the deposits of previous experience, and (b) the external stimulus situation. These largely correspond to the aggregated subjective desires and more external desiderata, which Eq. 35 seeks to itemize towards determining each objectively and towards approaching a synthesis which will predict behavior more accurately than our present rough and ready, all-or-none, observations are apt to do. *d. Von Wiese's formula for a process,  $P = A \times S$ .*

This process of tensing can be shown to correspond closely to Von Wiese's general formula for social actions (Ref. 78, pp. 174-176). Disclaiming mathematical exactitude, he suggests that "every social process (P in his notation) is the resultant ('product') of a personal attitude (A) and a situation (S)."

$$P = A \times S \quad (\text{Eq. 38a, Ch. X})$$

His "personal attitude" includes the intensity of desire, D, in the tension theory as one aspect of attitude; his "situation" includes the amount of the desideratum, V, as a part of the situation. The tension E, defined as the ratio of the intensity of desire of all the plurel observed to the desideratum,  $E = PDV^{-1}$ , is a ratio of a part of his concept of "attitude" to a part of his concept of "situation." The tension hypothesis comes in here to link the tension with societal action in the causal relation symbolized as:

$${}^tE, \bullet {}^u(I), ? = r > 0 \quad \text{the tension hypothesis} \quad (\text{Eq. 39, Ch. X})$$

This equation asserts the hypothesis (?) that a series of prior ( ${}^t$ ) tensings correlates positively ( $r > 0$ ) with an index of subsequent ( ${}^u$ ) societal actions (I) of some specified type ( ${}^|$ ). As these tensings are a ratio of measured desires and desiderata, they

cannot be expected to correlate perfectly with the societal actions of which they may be only a partial cause. The deficit from perfect correlation measures the remaining unknown causes of the societal action. In Von Wiese's formulation, however, his more inclusive but unmeasured "attitudes" and "situation" are set equal to the societal "process," thus connoting a perfect correlation between the attitude-situation product and the process. This seems a less realistic analysis of societal phenomena than the tension theory analysis in terms of measurable entities with proper allowance for their current imperfection.<sup>25</sup> \*

### 3. REINDICATING, $I^2$ , = SECONDARY INDICATORY PROCESSES

#### *a. Change of second moments.*

The primary indicatory processes, catalogued above, are definable as changes in the statistical first moment, the mean, including as its limiting case a change in a population of one, such as tensing in the case of one person. Next in the cataloguing come secondary indicatory processes. Here the aggregation of entities is always of more than one entity and the processes deal with internal movements within the aggregation. These are processes definable by change in the statistical second moment (see Eq. 9, Ch. V), from which fact these processes are here designated as "secondary" processes.

The second moment has two forms according as an indicant is multiplied by itself giving the variance, or by another indicant giving the covariance (i.e., the cross-product moment). As the variance (and its square root which is the standard deviation) measures deviations and the covariance measures correlation, changes in these will be termed the processes of "redeviating" and of "reкорrelating."<sup>26</sup> These processes may be measured by changes in a standard deviation or a correlation coefficient respectively.

$${}_t(I)_{I,I}^{.5} = {}_tI = \text{"redeviating," a sigma involving change in time} \\ \text{(Eq. 41, Ch. X)}^{27}$$

$${}_t({}_sI)_{I,J} = {}_tI = \text{"reкорrelating," a correlation involving change} \\ \text{(Eq. 42, Ch. X)}^{27} \dagger$$

\* For Eqs. 38b, 40a-b, Ch. X, see notes at end of chapter.

† For Eqs. 43a-b, 44a-b, Ch. X, see notes at end of chapter.

Each of these two processes subdivide into three subtypes which will next be described.

b. *Redeviating*,  ${}^{\circ}I$ , or  $I_{I.I.}^{\cdot 5}$ .

There are at least three forms of the redeviating process, depending on how the standard deviation is calculated, and each form describes a definite type of dynamic societal phenomena. These will be called "redispersing," "reordinating," and "re-varying."

(1) *Redispersing*,  $\pm {}^{\circ}I$ . "Redispersing" will be the label for those processes defined by a change in the standard deviation of some index in a population. Thus the standard deviation of the incomes of employees may be made smaller in a process of equalizing incomes, or they may be made larger in a process of dispersing incomes. A school class may be selected so as to be fairly homogeneous, thus equalizing ability, or may have pupils who are of very unequal ability added to it, thus dispersing the ability of the class.

A city may grow by adding millionaires at one extreme and slum dwellers at the other, thus increasing the sigma of social status in a process of dispersing. A nation may abolish its titled nobility in a democratic revolution, thus decreasing the sigma of aristocratic status in a process of equalizing. If the spread between the rich and the poor is decreasing, as alleged under pure communism, equalizing is going on, while if the spread is increasing, as alleged under unrestricted capitalism, dispersing is going on.

As neither the size of the population nor the mean wealth need be changing, there may be neither nullary nor primary processes going on.

Redispersing indicates the change in the spread as, for example, from the highest ranks of an army or any organization to its lowest ranks. It expresses the changing degrees of hierarchy, or subordination-superordination.

Whatever the content, or societal characteristics which are dispersed or equalized, the form of the process specified by a changing sigma is the same in the above illustrations, and this is taken to define the process:

$$\text{Redispersing} = \begin{cases} +_t(^{\sigma}I) = \text{"dispersing," the increase of the} \\ \quad \text{sigma of a characteristic} & (\text{Eq. 45a, Ch. X}) \\ -_t(^{\sigma}I) = \text{"equalizing," the decrease of the} \\ \quad \text{sigma} & (\text{Eq. 45b, Ch. X})^{28} \\ & (\text{See S. 55, Ch. X}) \end{cases}$$

As with other processes, the definition is based on the statistical form of the S-equation and applied to any societal content. Redispersing may be of income or intelligence, prestige or avoidrdu-pois, or of any economic, political, religious, or other societal characteristic, in whatever I-units expressed.<sup>29</sup>

A process would usually have its content specified or understood from context, as in "dispersing of income," "equalizing social status," etc. This definition by identical form of formula in no way implies identical relations of all dispersed "contents" to other phenomena. Thus equalizing income need not correlate with the political party in power simply because one party tends to equalize the tariff on commodities, while the other party does the reverse. Each "content" will have to be studied alone, just as Physics studies the velocities of different objects in relation to their other properties separately for each object. But the process is an operational definition of a common form of societal action. Given the formula, Eq. 45, different operators can calculate the sigma of given data and obtain identically the same thing defined—an amount of change in dispersion.

(2) *Reordinating*,  $^{\sigma}(_tI)$ . The next subprocess of redeviating is "reordinating," defined as the sigma of changes of a characteristic in a population in one period:

$$(_tI)_{I-1}^{\cdot 5} = ^{\sigma}(_tI) = \text{"reordinating," the sigma of changes (of one} \\ \text{characteristic between two dates in one popula-} \\ \text{tion)} & (\text{Eq. 46a, Ch. X})^{30} * \\ & (\text{See S. 36, 39, 47, Ch. X})$$

This is the sigma of a difference in time.<sup>31</sup>

The most important subform of reordinating is when the index is a desideratum, as this defines "competing," the process of competition.

\* For Eqs. 45c and 46b, Ch. X, see notes at end of chapter.

(a) *Competing* (or “*regrading*”), *Cp*. The outstanding characteristic of competing is that each party wants to obtain a larger share of the limited desideratum, *V*, for which they compete. By as much as one competitor wins the other loses. Each customer gained by *P* is one less customer, whether actual or potential, for *Q*. Thus, Japan’s textile trade gains in South America are at the expense of other competitors. The victory of one football team is the defeat of the other. Each competitor strives for exclusive possession, or at least a larger share, of the limited desideratum, thereby displacing other competitors. This shift of some of the quantity of the desideratum from some competitors to others offers a way to measure, and so to define, competing. Let us then define *effective competing* during a period as the transferring of the desideratum from some competitors to others. The standard deviation of these gains and losses seems the best measure to summarize the extent of such transfers. (Note that the mean transfer is zero, as gains must equal losses by definition.) Any net gain or loss for the whole population measures the “progressing” or “regressing” process, as defined above, that has taken place in addition to the competing process.

To illustrate the measurement of the process of effective competing consider the simple case of two competitors. In 1820, 90% of United States foreign trade was carried in United States vessels, but only 10% was so carried in 1900, the rest being carried in foreign vessels, almost all British.

	<i>Beginning of Period (1820)</i>	<i>End of Period (1900)</i>	<i>Gain or Loss</i>	<i>Square of Difference</i>
First competitor’s <i>V</i> . . . . .	90%	10%	– 80	1600
Second competitor’s <i>V</i> . . . . .	10%	90%	+ 80	1600
			0	3200

$$\text{Mean change} = 0 \quad \sigma_v = \sqrt{\frac{3200}{2}} = 80\% = \text{percentage of maximal competing}$$

The index of effective competing between the American and foreign marines in these eighty years was 80%.<sup>32</sup> This is 80% of the maximum possible intensity of the process of competing such as would exist when the initial monopoly of one competitor became the terminal monopoly of the other, yielding a  $\sigma$  of 100%. The minimum competing at the other limit would be 0%, as when all competitors end up with the same relative shares they

started with, so that no transfer of the desideratum would have taken place.

This formula is general to any number of competitors, P, and a desideratum in any units which can be re-expressed as each competitor's percentages of the total desideratum. Let V denote a competitor's percentage. Let  $\iota V$  denote any change in that percentage, whether gain or loss, in the period T. Then the index of effective competing is:

$$C_p = (.5 \iota V_{\Sigma p}^2)^{.5} = \text{the index of effective competing}$$

(Eq. 47a,\* Ch. X)

(See S. 10, 24, 40, 42, and 52, Ch. X)

The summation here is either over the  $|_p$  plurals or the  $|_p$  persons, depending on which are the competitors in a specified situation. This index is a percentage measure of the actual amount of shifting of the desideratum between competitors. The amount is taken as a percentage of the maximum possible amount of shifting. It is derived as a ratio of the standard deviation of the gains and losses of the competitors to the standard deviation of the situation where the initial monopoly of one competitor becomes the terminal monopoly of another.<sup>33</sup> †

"Competing" as defined above is "effective competing," and must be distinguished from "causative competing." It is measured (as are all the processes in this chapter) by its effect. The intensity of competing is measured by the gains and losses that result. "Causative competing" is the effort put forth by competitors—which may or may not result in any effective competing—for intense efforts of competitors may cancel each other and result in no gains or losses, as often happens in advertising. Causative competing, though highly important, is less tangible and requires inventing special indices, such as the cost and area of advertising space, etc. Effective competing has the immense advantage for scientific purposes of being more readily measured. But the reader should clearly note that the effective competing (as with all the "effective" processes defined here) is only part of the phenomena of competition. In terms of the basic tension equation, Eq. 34, Ch. V (or Eq. 35, Ch. X), the causative competing would be reflected by E, the tension, for the tension in-

\* For Eqs. 47b-47d, see note 34 at end of chapter.

† For Eqs. 47b-52, Ch. X, see notes at end of chapter.

creases, the competitive effort needed becomes greater, when the number of competitors increases, or their desires go up, or the desideratum becomes relatively scarcer. But the effective competing is defined as the actual *redistributing* of a given amount of a desideratum between the competitors. The possible amount of redistributing decreases as the number of competitors increases. As defined by these formulae, it can be proved that effective competing varies inversely with the number of competitors, while causative competing or tensing varies directly with the number of competitors.

It should thus be evident that our competitive capitalistic economic system is only partially competitive. While each competitor is struggling to obtain as large a share of the desideratum for himself as he can, this causative competing (i.e., motivation or tension, E) makes him work to produce more of the desideratum, which is our definition of progress in its subform of effective co-operating,  $+_tV_{co}$ . In the production of wealth (and other desiderata), its redistribution is a by-product, and a by-product controlled in varying degrees by governmental action to prevent its becoming monopolistic. The formulae proposed here enable the isolation of the competitive from the co-operative component in any situation where the desiderata involved are measurable (including the case of unitary qualitative desiderata). The co-operative component is measurable by the mean amount of change in the desideratum, the statistical first moment of changes ( $= \Sigma_t V/P = M_v$ ), while the competitive component is measurable by the redistribution of the desideratum, the statistical second moment of changes ( $\Sigma_t V^2/P = \sigma_v^2$ ).

(b) *Revaluating*. Reordinating processes may next be illustrated in the case where the index measures the intensity of desire of the persons, or of the plurels, in a population. This will be a sigma of changes in D in the tension theory (Eq. 34, Ch. V). For a simplified numerical illustration, consider the shifts of public opinion in the United States as measured by one of the successive polls before the presidential election of 1936. (Ref. 33.)

	Votes for Roosevelt	Votes for Other Candidates	
January, 1936.....	60.8%	39.2%	
October, 1936.....	59.2%	40.8%	
Gains or losses = $_tD$ =	-1.6%	+1.6%	RI = 1.6%

According to these data the election campaign hardly affected the desires of the voters, as the net redistribution of valuation as measured by our index of "revaluating" is only 1.6%. This index of "revaluating" has the same form of formula as before, with the intensity of desire ( $\nu D$ ) substituted for the desideratum ( $\nu V$ ) and for the population ( ${}_{\infty}\nu P$ ) in the formula for net mobility. As before, the revaluating index varies from zero to a maximum of 100%. It is the ratio (in percentage units) of the observed sigma to the maximum sigma. The maximum sigma would occur when all the population's desire was initially concentrated on one desideratum and was finally transferred completely to another desideratum.

$$R1 = (.5\Sigma\nu D^2)^{.5} = \text{index of revaluating} \quad (\text{Eq. 53a,* Ch. X})$$

(See S. 10 and 18, Ch. X)

The election illustration given above happens to be a versatile one, as it illustrates secondary processes of all the four factors in the tension equation (Eq. 34, Ch. V). The votes are a desideratum competed for by the Roosevelt and anti-Roosevelt parties; they also represent the net mobility of persons between the pro-Roosevelt to the anti-Roosevelt plurels; they also measure the shift in popularity, or relative intensities of desiring, between two desiderata by the electorate; and finally, since in this situation the societal tension equals the average desire, the shift of votes measures the relative "retensing," or shift of tension, of the electorate towards the two camps. The votes happen to measure these four factors simultaneously in this situation, because the unit of desire in a democracy is defined as one vote and is identical with one person. The votes are the *desideratum* (V) competed for *to the candidates*; *to the electorate* each candidate is the unitary qualitative value, the intensity of *desire* (D) for whom is measured by the votes.

Whether some quantity measures the quantity of a desideratum or the intensity of desire for it depends on the plurel relative to which the desideratum is defined. What is a desideratum to one plurel may not be such to another. The desideratum and the plurel and the desire can only be defined relatively to each other.

Revaluating is here illustrated as between several desiderata;

\* For Eq. 53b, Ch. X, see note 34 at end of chapter.

but revaluating may also exist between several plurels with respect to one desideratum. Thus on *revaluating between desiderata*, the  $\Sigma$  in Eq. 53, Ch. X is over the  $v$  desiderata; while in *revaluating between plurels*, the  $\Sigma$  in Eq. 53, Ch. X would be over the  $p$  plurels.<sup>34</sup> \* The former expresses the relative shifting of popularity of various desiderata to one plurel; the latter expresses the relative shifting of popularity among various plurels with respect to one desideratum.

(3) *Revarying, Rv.* The third type of redeviating process, after redispersing and reordinating, remains to be described. It is the process measured by the standard deviation of one characteristic observed in one population on a series of dates. If this characteristic is constant through the series of dates, its sigma will be zero and its stability will be maximal (for that period, of course). If the characteristic fluctuates in amount from date to date, its sigma measures this tendency to revary.<sup>35</sup> †

This sigma varies from the lower limit of zero, where stability is greatest, upwards without any general upper limit as stability decreases. Thus, if the annual income of an institution fluctuates very little from year to year over a period of years, we say that that institution has financial stability. If the language of a people (such as the Turks under Kemal Ataturk) is officially dropping a considerable percentage of its words (because of Arabic and non-Turkish roots) in a long black-list every year and is adopting new words (based on Turkish roots), then that language during this period has been unstable. (About fifty percent of the vocabulary of the Turkish language in the official dictionaries is alleged to have changed in the two decades since the World War.)

Heredity may be defined as a zero amount of revarying between generations. That trait is most purely hereditary which varies least in a series of generations. Traits that fluctuate in amount between generations are less purely hereditary, i.e., less stable. The stability may be due to the biological heredity in the germ plasm (preconception factors) or to cultural heredity of training and other environmental influences (postconception factors). Thus the Chinese skin color is biologically stable, while the Chinese language for many generations past has been culturally

\* For Eqs. 54a-57e, Ch. X, see notes at end of chapter.

† For Eq. 58a, Ch. X, see notes at end of chapter.

stable. Revarying is measurable in degrees and thus redefines heredity in more operational terms. Of course, it merely measures the degree of revarying, reflecting the gross influence of heredity, and does not analyze the mechanism of transmission from parents to offspring.

$Rv_H = {}^{\sigma}({}^{t:p}I)$  = revarying between generations, a measure of heredity (Eq. 58b, Ch. X)

This Brief-S formula directs the operator to calculate the standard deviation of an index observed on a series of dates (generation-time intervals,  ${}^t|$ ) to each of which there corresponds a number of persons (generation-plurels,  ${}^p|$ ). This is a measure of the degree to which a given trait may be considered to be hereditary. (Of course, another equivalent index may be more suitable to a given situation than a sigma to measure stability, depending on the nature of the data and the investigator's purpose).<sup>36</sup>

(c) *Recorrelating*,  ${}_{\sigma}I_I \cdot J$ . The three redeviating processes of redispersing, reordinating, and revarying are subforms of the second self-moment, or variance, entered in the main diagonal cells of the matrix Eq. 43, Ch. X. The "recorrelating" processes defined by the non-diagonal cell entries of Eq. 43 are to be studied next. These are the covariances, or cross-product moments. In vectorial algebra they are the scalar products of two different vectors.

There are four subforms of the recorrelating process (Eq. 42, Ch. X) which will be designated "recodispersing," "reco-ordinating," "reranking," and "recovarying."

(1) *Recodispersing*,  ${}_{\sigma}I_I \cdot J$ . Parallel to redispersing, which is the change of a sigma, is recodispersing, which is the change in time of a correlation coefficient of some kind. Both are changes in the second moment.<sup>37</sup> \*

Thus in S. 48, Ch. X the split-half reliability correlation of each of a set of intelligence tests was observed on one date and reobserved on another date at a second trial on those tests by the same population. Practice factors and longer periods for work tend here to increase the reliability correlations, to make the recodispersing positive.<sup>38</sup> †

Recodispersing answers the questions, "Are two different quantitative characteristics becoming more or less qualitatively alike

\* For Eq. 59, Ch. X, see notes at end of chapter.

† For Eq. 60, Ch. X, see notes at end of chapter.

as time passes?" and "How much so?" In geometric terms where the size of an angle represents the degree of qualitative similarity, recodispersing measures the amount of change of the angle. Thus in S. 34, Ch. II the prediction of graduation standing from early academic record shows positive recodispersing as the academic record grows longer and graduation comes nearer. This is graphed as progressively smaller angles between the graduation vector and the vectors of progressively cumulating academic record. (See Chapter XII for fuller development of this principle that prediction correlations generally increase with immediacy of the events predicted.)

(2) *Reco-ordinating*,  ${}_c(\nu I)_{I \cdot J}$ . In addition to recodispersing, which is the change of a correlation, there is the correlation of changes which will be labeled "reco-ordinating." <sup>39</sup> \*

Changes may go on with different velocities as between different characteristics, or as between different plurels. This varies the pattern of their relationships. It alters their co-ordination and hence the term "reco-ordinating." These phenomena in simplest forms are better known in the sociological literature by such terms as, "culture lead and lag." But hitherto "lead and lag" have implied comparison of traits in pairs, whereas the correlation coefficient provides a summarizing measure for changes in more than a pair of traits, and thus reflects a changing *pattern* of culture traits more adequately.

Consider first the simplest forms of culture lead and lag. This phrase means that some traits have developed faster than others and are observed to be ahead of those others, thus giving rise to strains and maladjustments in the observed culture. This is the dominant point of view of such a comprehensive study of change on a national scale as *Recent Social Trends*.<sup>40</sup> (Ref. 56.) The concepts of lead and lag imply that, beginning with a state of working equilibrium, or internally consistent culture, there follow changes of culture traits at differential velocities. Thus out of a working balance between government organization and transportation in the county system in the United States of a century ago, the more rapid development of the auto and other transportation in recent years has created a maladjustment because county government organization has lagged behind, chang-

\* For Eqs. 61a-b, Ch. X, see notes at end of chapter.

ing little, if at all. Since to decelerate the development of transportation is undesirable, the alternative is to accelerate change in the county government by associating counties together into larger units with consequent reduced overhead and taxation. Similarly the doctrine of national sovereignty worked satisfactorily in days of states with little interdependence. But the rapid growth of transportation and communication, and economic and other interacting between nations has outstripped the slower development of effective international regulation.

Such differential velocities of cultural changes may be symbolized by the velocities of an aggregation of indices of culture traits and complexes as follows:

(I)<sub>i/t</sub> = inter-index lead and lag, or differential velocities of changes (Eq. 62a, Ch. X) <sup>41</sup> \*

But the concepts of lead and lag apply between plurels as well as between indicators. Thus in any given indicator of culture some peoples will be "advanced" and others "backward." The illiterate masses of Asia and Africa lag behind the Scandinavian countries in literacy. The West led the East in industrializing. Japan "modernized" more rapidly than China in the last century. Urban culture changes more rapidly than rural culture. Generalizations such as these can be more exactly formulated in formulae which require exact specification of their qualitative and quantitative ingredients, such as:

(I)<sub>p/t</sub> = T<sup>-1</sup> :  $\underline{P}_p$  : (I) = inter-plurel lead and lag  
 (in Brief-S notation) (in full notation) (Eq. 62c, Ch. X)

This asserts the velocity of change of some specified index corresponding to each of  $\lfloor_p$  plurels. With this formula as a basis the difference between any two plurels and its reliability <sup>42</sup> † may be stated and a host of relations can be quantitatively explored. Issues of invention versus diffusion of culture can be more precisely investigated.

Next consider lead and lag in more complicated patterns than a comparison of the two members of a pair. Suppose we wish to measure the relationship, the co-ordination of changes, in two dynamic characteristics such as birth rates and death rates in a

\* For Eq. 62b, Ch. X, see notes at end of chapter.

† For Eq. 63, Ch. X, see notes at end of chapter.

series of plurels such as the 48 states of the United States. The correlation coefficient of such a pair of rates of change in a series of plurels measures the process of reco-ordinating (of two characteristics in one population of  $\frac{1}{p}$  plurels). Another pattern might be to obtain the correlation between Protestant birth rates and Catholic birth rates in the 48 states in order to isolate religious and geographic components of birth rate phenomena. This correlation measures the process of reco-ordinating (of one characteristic in two populations of  $\frac{1}{p}$  plurels).

Further variants of reco-ordinating are possible. But enough has been suggested to make the student of Sociology aware that culture lead and lag may be analyzed in many different formal ways, of which three general types are:

<i>Change</i>	<i>Measuring Index</i>
1. in a pair of characteristics in one plurel in a pair of plurels in one characteristic	by a difference of indices (and its significance ratio)
2. in many characteristics in one plurel in many plurels on one characteristic	by reordinating ${}^c(tI_i \text{ or } p)$ and its standard error) <sup>43</sup>
3. in two characteristics in many plurels in two plurels in many subplurels	by reco-ordinating $(tI_p)_{I \bullet J \text{ or } p' \bullet p''}$ (and its standard error)

All of these six types are ordinating processes of various degrees of complexity and in various arrangements. They can measure the changing of the observed co-ordinations in a given culture and period.

(3) *Reranking*, " $\bullet\bullet\bullet$ " $I$ . A part of the reordinating process is reranking of the entities in a frequency distribution (the remaining part being a function of redispersing as shown exactly in Eq. 46b, Ch. X). In competing, the competitors become reranked (in respect to the amount of the desideratum competed for that each acquires), and the dispersion of the competitors may also change. In mobility, the plurels become reranked as to the size of their membership as well as becoming redispersed. This reranking process can be isolated by itself and measured as the

correlation between an index in a population on one date and that index on a later date.<sup>44</sup> \*

A reliability correlation determined by repetition of the measurement on two occasions is an example of reranking.<sup>45</sup>

Repetition of a race resulting in a changed ranking of the competitors is an instance of the reranking process. Shifting social status of occupational plurels is another example of reranking (though if the dispersion of status as measured on some scale also changes, the more inclusive process of reordinating has gone on). Shifting popularity of candidates for an election, especially if expressed in rank order, is another instance of reranking.

(4) *Recovarying*,  $({}^tI)_{I \cdot J}$ , *Correlation of Time Series*. The fourth type of correlating process parallels revarying. It is the familiar correlation of two time series. Do the prices of commodity I tend to rise and fall with commodity K? Does the time curve of phenomenon I tend to parallel the time curve of phenomenon J? The tendency to align two time curves can be measured by a correlation coefficient which here is taken as defining a process, which for convenient identification will be labeled "recovarying":

$({}^tI)_{I \cdot J}$  = "recovarying," the correlation of two time series

(Eq. 65, Ch. X)

(See S. 27 and 28, Ch. X)

This Brief-S formula asserts a correlation between two different indices,  $|_I$  and  $|_J$  (each expressed in sigma units), observed on a series of dates.<sup>46</sup> All the time curves in the situations appended to this chapter are examples of this recovarying process. (Unless their correlation is explicitly calculated, they are analyzed in the descriptive and quantic formula as aggregates of changing indices ( $T^{-1} : I_i$ ), however, and not as correlated indices ( $T^{-1}I^2$ ). This is because the function of the quantic formula is to state the operational degree to which the relationships within the situation have been explicitly developed.)

In concluding this exploration of secondary indicatory processes<sup>47</sup> it may be noted that there are further types which have not been described here, as they are still of infrequent occurrence in the sociological literature.

The indices defining processes (which are particular forms of

\* For Eqs. 64a-64b, Ch. X, see notes at end of chapter.

change in the total flux of societal on-going) are attempts towards standardized analysis of such societal patterns. Our standardized secondary indices analyzing certain aspects in defined situations of the total pattern of societal on-going may be summarized in the table below:

### REINDICATING = $\frac{1}{2}I^2$ , THE SECONDARY INDICATORY PROCESSES

$${}_{\nu}(I)_{I \bullet I'} = T^{-1} : \Sigma II' / N = T^{-1} I^2 = 9;2;1;p$$

a scalar product, Brief-S notation      a second moment with change, Full-S notation      change in an index to the second power, the quantic formula      the quantic number

$N = v, i, p, P$

$\sigma$	$r$
<p><i>Redeviating</i> = <math>{}_{\nu}(I)_{I \bullet I} = {}^{\sigma}I</math></p> <p>defined by change in (the square root of) a variance, a function of the self-moment, <math> _{I=I'}</math></p>	<p><i>Recorrelating</i> = <math>{}_{\nu}({}_{\sigma}I)_{I \bullet J} = {}^rIJ</math></p> <p>defined by change in the covariance (in sigma units), a function of the cross-product moment, <math> _{I=I'}</math></p>
<p>a. <i>Redispersing</i> = <math>{}_{\nu}(I_{I \bullet I})^{\frac{1}{2}} = {}_{\nu}({}_{\sigma}I)</math></p> <p>Dispersing = + (); Equalizing = - () Change of a <math>\sigma</math></p>	<p>a. <i>Recodispersing</i> = <math>{}_{\nu}({}_{\sigma}I_{I \bullet J}) = {}_{\nu}({}^rIJ)</math></p> <p>Change of an <math>r</math></p>
<p>b. <i>Reordinating</i> = <math>({}_{\nu}I)_{I \bullet I}^{\frac{1}{2}} = {}^{\sigma}({}_{\nu}I)</math></p> <p>A <math>\sigma</math> of changes, the dispersion of one dynamic index</p>	<p>b. <i>Reco-ordinating</i> = <math>({}_{\sigma\nu}I)_{I \bullet J} = {}^r({}_{\nu}IJ)</math></p> <p>An <math>r</math> of changes, the correlation of 2 dynamic indices</p>
<p>c. <i>Revarying</i> = <math>({}^tI)_{I \bullet I}^{\frac{1}{2}} = {}^{\sigma}({}^tI)</math></p> <p>A <math>\sigma</math> of a time series, i.e., the dispersion of an index on a series of dates</p>	<p>b'. <i>Reranking</i> = <math>({}^{\nu'} \bullet {}^{\nu'} I)</math></p> <p>The correlation of an index at one time with itself at another time</p> <p>c. <i>Recovarying</i> = <math>({}^tI)_{I \bullet J} = {}^r({}^tIJ)</math></p> <p>A correlation of 2 time series</p>

The first formula after each process name is in Brief-S notation; the second one is more conventional notation using the symbols  $\sigma$  and  $r$ . The first formula shows the calculational operations more explicitly and shows the parallelism of the redeviating and the recorrelating indices.

Note that "ordinating" is a function of redispersing and reranking:

Reordinating is calculable by the formula for a sigma of a difference (in time).

Reco-ordinating is calculable by the formula for a correlation of two differences (in time).

### 4. SKEWING, $\frac{1}{3}I^3$ , = TERTIARY INDICATORY PROCESSES

Tertiary indicatory processes are defined as changes measured by the third statistical moment. The most common type is a

skewing which is a change in skewness and may be defined by a change in Sk of Eq. 10, Ch. V:

$$T^{-1} : \Sigma_e I^3 P^{-1} = {}_tSk = \text{skewing} \quad (\text{Eq. 66a, Ch. X})^{48}$$

When the skewed index measures a desideratum, V, then negative skewing measures the process of "ameliorating."

$-{}_tSk_v =$  "ameliorating," negative skewing, or reducing positive skewness of desiderata

(Eq. 66b, Ch. X; cf. Eq. 29, Ch. V)

Thus when a government's policy is towards redistribution of wealth by imposing more steeply progressive inheritance and income taxes, and reducing sales taxes or other taxes whose incidence is greatest on subaverage incomes, in addition to redistributing a negative skewing process is going on, tending to reduce the positive skewness that normally exists in incomes. (See S. 20, Ch. XII.)

This negative skewing of a distribution of a desideratum is "ameliorating." It is reducing the socially undesirable extreme positive skewness (as of incomes). Wherever positive skewness is a negative desideratum to a population, there negative skewing is a positive desideratum and may be termed "ameliorating." (See Chapter V for a further suggestion concerning amelioration.)

Higher order processes measured by changes in the fourth moment (change of kurtosis) and by higher moments can be defined when needed.

## 5. SUMMARY OF THE INDICATORY PROCESSES, ${}_tI^i =$ INDEXING

The large field of the indicatory processes may now be reviewed. These processes are what are measured by changes in indicators. An indicator is a human record of any characteristic of people or of their environment, other than time, space, and the number of people. Indicatory processes were classified by their exponents as shown in the successive statistical moments (Eqs. 7-11, Ch. V) of frequency distributions. These moments reflect the operational degree to which the data have been developed.

Nullary processes deal with change in a *quality* and in the number of qualities. Primary processes deal with change in a *quantity* and in the mean quantity. Secondary processes deal with change in the internal *relations* of the characteristics of a population—

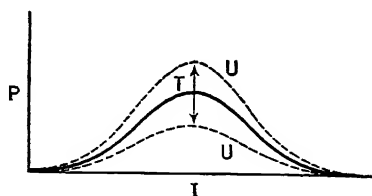
either deviating on either side from a reference point, or correlating, which is deviating on either side of a reference line. Tertiary processes deal further with internal relations, especially the symmetry of the deviating.

These processes may be graphed as forms of movement of a multitude of societal entities which are the units of frequency (or the "distributed" units—cf. Chapter V) of a frequency distribution. These entities are usually persons or plurels (but may be dates or other indices or descripts).

#### DIAGRAMS OF PROCESSES DEFINED BY STATISTICAL MOMENTS

##### I. Nullary processes—zero-order moments, $T^{-1} : \Sigma I^0 P^{-1}$

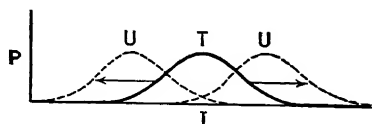
Change in *frequency*



The area under the curve changes, i.e., the population increases or decreases

##### II. Primary processes—first-order moments, $T^{-1} : \Sigma I^1 P^{-1}$

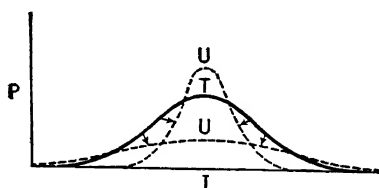
Change in *mean*



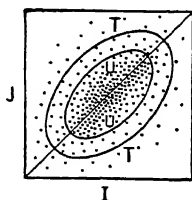
The whole curve moves in one direction, i.e., the population increases or decreases its mean characteristic

##### III. Secondary processes—second-order moments,

A. Change in *sigma*  $T^{-1} : \Sigma_d I^2 P^{-1}$



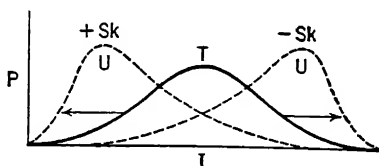
The curve moves in two directions from the center, i.e., the population is dispersing from, or concentrating towards, the mean,

B. Change in correlation  $T^{-1} : \sum_{\sigma} I_{\sigma} J P^{-1}$ 

The scatter moves in two directions from the reference diagonal, i.e., the population is dispersing from, or contracting towards, the diagonal line

IV. Tertiary processes—third-order moments,  $T^{-1} : \sum_{\sigma} I^3 P^{-1}$ 

Change in skewness



The curve moves asymmetrically in one direction, i.e., the population increases on one side of the center

Key:

Arrows show the direction of movement of the population,  $P$ , with respect to the indicator,  $I$ , of the distributing characteristic.

$t'$  = curve at the time when the process starts

$t''$  = curve later in the process

In IIIA Redeviating processes:

Redispersing = change from  $t'$  to  $t''$

Reordinating = curve  $t'$  where  $I$  is  $t' - t'I$

Revarying = curve  $t'$ , where  $P$  is replaced by dates,  $t'T$

In IIIB Recorrelating processes:

Recodispersing = change from  $t'$  to  $t''$ , or  $t''$  to  $t'$

Reco-ordinating = where  $I$  and  $J$  are dynamic,  $t'I$  and  $t'J$

Reranking = scatter  $t'$ , when  $J$  is  $t'I$

Recovarying = scatter  $t'$ , where  $P$  is replaced by  $t'T$

(Normality of curves is not assumed)

In addition to the processes defined by a summarizing moment some changes of aggregates before they have been so summarized were discussed.

All this means that, within a sector, the classification is based in large part on the statistical *form* of the societal change rather than on the societal *content* of the change. Processes are defined by frequencies, means, sigmas, and correlations, not by religious,

educational, economic, and other types of content. This principle is a corollary of the definition of Sociology stated in this volume, as the study of the "general characteristics common to all classes of societal phenomena." The content of change is more specific to the field of changing data, while the forms of change, according to the working hypothesis, are common to all the social sciences. In studying forms of change, those forms were chosen which can be operationally defined, for these can be communicated and duplicated and verified—essential requirements in building a science. The statistical forms, such as the formulae of this chapter, are operational definitions,<sup>49</sup> since they tell the operator what calculations he must perform to obtain what the formula defines. Any two competent operators, given the same recorded data, will then obtain the same means, sigmas, etc., and hence reach objective and complete agreement as to what the processes are they are studying.

But the forms common to all societal phenomena are not exclusively statistical. Thus intensity of desire (D) and desiderata (V) and resulting psychological tensions (E) are highly general forms or characteristics in all the social sciences, and, as a start, these psychological forms have been combined with the statistical forms. The result has been that combinations and permutations of these forms, expressed in manipulating the scripts and operational symbols of our basic S-equation (Chapter II), have yielded all the processes described in this chapter and many more. This theoretical formulation was found to fit all the 1500 quantitatively recorded situations of societal change which have been found and analyzed to date, and also to fit a large body of current sociological concepts. Thus current concepts such as interaction, isolation, contact, mobility, conflict, accommodation, co-operation, competition, differentiation, association, progress, tension, and others in the rest of this chapter can be more exactly defined in operational terms by using the formulae of S-theory.

It is expected that in each social science and subfield of it specific content may be combined with these general forms yielding subprocesses appropriate to that field. Thus the content of economic desiderata fitted into the form of "reordinating" yields an operational definition of "economic competing."

*C. Spatial Processes,  $L^1$  = "Spacing"*

Changes in populations and in indicators have been explored. Changes in physical space which are of societal reference will next be systematically explored. These spatial changes have a greater variety of forms owing to the four exponents of spatial indices. Since, however, these spatial processes are of less sociological significance than populational and indicatory processes, they will be discussed more summarily.

Spatial processes may be classified on the basis of the operation of the exponent into four classes of "spatial dimensioning," as changes of points, lines, areas, or volumes.<sup>50</sup> \* To be societal processes as distinct from geometric or other non-societal processes, the points, lines, areas, and volumes must be qualified, of course, by qualities of societal significance, as in identifying the points as homes in a city (S. 13, Ch. VIII), centers of population in a country (S. 58, Ch. X), etc.; or the lines as traffic routes (S. 23, Ch. VIII; S. 68, Ch. X), etc.; or the areas as neighborhoods (S. 4, Ch. VIII), zones with defined societal characteristics (S. 22, Ch. VIII), nations (S. 61, Ch. X), etc.; or the volumes, as water used per capita (S. 73, Ch. X), cranial capacity (S. 12, Ch. IX), etc.<sup>51</sup>

In the case of the space sector, unlike the other three sectors, any spatial index of change may have any one of the four exponents and may then be further observed in a frequency distribution. Societal situations often involve a collection of points, a collection of lines or areas, etc. As usual, the statistical moments of these frequency distributions may define further forms of change, which are called the processes of "spatial distributing." Thus there are changes in the total frequency of the distribution, i.e., changes in the number of points, or in the number of lines, or in the number of areas. There are changes in the mean length of lines and in the mean size of areas and volumes. There are changes in the dispersion of lines, areas, and volumes. These are all defined by the moments representing a second operation of exponents upon the spatial indices.

This double classification of spatial processes by two successive operations of the exponent may be clarified by a matrix cross-classifying them.<sup>52</sup>

\* For Eqs. 67a-d, Ch. X, see notes at end of chapter.

In addition to the processes defined by these two successive operations of the exponents upon spatial indices, there are further changes of these indices in combination with other indices, P, I, T, and L again.<sup>53</sup>

Among these processes, populational densating, Eq. 67cP, is of greatest sociological significance. In the form of the ratio of persons-per-area it is labeled the "densing" process. It is the change of the man-land ratio, or the dynamic aspect of populational density.

$$\text{t}(PL^{-2}) = \text{"densing"}^{54} \quad (\text{Eq. 68, Ch. X}) \\ (= \text{Eq. 67c1 P/L})$$

Densing may measure such current terms as: ruralizing, urbanizing and suburbanizing, settling, colonizing, depopulating, migrating, commuting-into-town, meeting, crowding, slum-clearing, housing, succession, invasion, and many other ecological concepts. These subforms are definable by the population, the area, and the period, involved in Eq. 68, Ch. X.<sup>55</sup> Closely correlated are such processes as industrializing, the prior changes from nomadic to settled, and from predominantly agricultural to commercial cultures.

One other spatial process may be singled out for special comment. It is the reordinating form of rearealing, defined by a sigma of changes of areas. Thus when the map of Europe is redrawn by a Versailles Treaty, the sigma of the percentage gains and losses of area among the nations (including the newly created nations) is a remapping process with a formula<sup>56</sup> \* identical in derivation to the formulae for competing, mobility, revaluating, etc. This formula expresses the degree of redistribution of territory as a percentage of maximal redistribution, as would occur if an initial monopoly of the area of Europe held by one nation changed to a terminal monopoly by another nation.

#### *D. Durational Processes, $\text{tT}^{\text{t}}$ = "Timing"*

The simple processes involving the indicatory, populational, and spatial sectors have been outlined. There remain the processes involving durations—the change, during periods of one kind, of the length of periods of another kind. (The kind of period is

\* For Eq. 69, Ch. X, see notes at end of chapter.

specified as usual by the implicit attribute-time product, i.e., the condensing of the attribute's descripts identifying a quality into the temporal descripts on the time index. See Eq. 8, Ch. IX.)

Only the two classes of durational processes defined by zero-order and first-order moments have been observed as yet in the literature of the social sciences, though second-order moments and other classes can be imaginatively described. These three classes are described by temporal exponents of 0, +1, and +2, and may be labeled "periodizing," "durlating," and "redurlating," respectively.

### 1. PERIODIZING, $\iota^{\cdot}(\Sigma^{\cdot}T^0)$ , = NULLARY TEMPORAL PROCESSES

Events of a given kind may occur at a slow or fast tempo, few or many in a given period. The number of events, dates, or punctiform temporal phenomena, expressed by the date script, measures a process which may be labeled as "periodizing." To "periodize" is defined by Webster as "to end." Every event or date ends a time interval since the previous event or date. Periodizing, then, is ending time intervals and is measurable by counting the events of the type studied which occur in the total period that is observed. Thus the frequency of occasions in S. 52, Ch. X, on which Moreno's subjects of "spontaneity tests" expressed anger, fear, etc., towards a given person, is an example of such periodizing. This process is positive only—the number of dates can only increase and can never sink below zero.<sup>57</sup> \*

Differences in the number of events, i.e., in the amount of periodizing in each of two different periods, can be negative as well as positive, but this is a change in the velocity of the periodizing process and is, therefore, of the order of an acceleration.

The acceleration of this process is an important index of social forces. Thus in a situation where a government is combating lawlessness and guerilla warfare, the number of incidents (such as raids with casualties per period) measures from period to period whether the forces of law and order are gaining or losing.

### 2. DURLATING, $\iota^{\cdot}T^{+1}$ , = PRIMARY TEMPORAL PROCESSES

Durlating is the changing of a simple duration of some kind. Its simplest form is a changing attribute-duration product. (See

\* For Eqs. 70a-b, Ch. X, see notes at end of chapter.

S. 83 and S. 85B, Ch. X.) Usually, however, a population is involved, implicitly or explicitly, in the situation. The commonest forms of durating are a changing life expectancy, or changing hours of work, study, or other use of time.<sup>58</sup> \*

### 3. REDURATING, $\frac{1}{2}T\frac{1}{2}$ , = SECONDARY TEMPORAL PROCESSES

All the secondary processes of redeviating and recorreling and their subprocesses may occur among durations. Thus when the length of the working week is made more equal for all classes of employees, the durational equalizing process,  $-_t(^{\sigma}T)$ , is going on, while if unemployment is making for greater inequality of duration of employment, the durational dispersing process,  $+_t(^{\sigma}T)$ , is taking place. Durational revarying is measured by the sigma of durations, such as hours of sleep, observed on many dates. When a timetable, schedule, or calendar, of past or future events is changed, this durational reordinating process may be termed "rescheduling." The standard deviation of the time differences between the new and the old dates measures the amount of rescheduling, just as competing, mobility, and other reordinating processes are measured. Rescheduling applies not only to the redistribution of time intervals within one overall period, but also to the redistribution of durations of any kind between a set of persons or plurels or other entities. Thus revising a school curriculum, or reassigning the periods of use of any public institution to various groups of patrons, may also be measured as a rescheduling process.<sup>59</sup> †

Many a student will naturally ask, "What is the use of such detailed formal analysis of processes, in a situation in which a glance at a graph tells its story?" The answer is both theoretical and practical. From a practical viewpoint, executives require averages and other summarizing indices in order to maintain more exact control, especially as phenomena in their field become more complex and organized on a larger scale. The process names and indices defining them are standardized contributions to this need for the entire field of dynamic societal phenomena. From the viewpoint of theory, two points may be noted—the advantages of a standardized classification and a realistic approach to societal analysis. Classification is an essential step in building a

\* For Eqs. 71a-c, Ch. X, see notes at end of chapter.

† For Eqs. 72-77, Ch. X, see notes at end of chapter.

science, in imposing conceptual order upon the chaos of raw phenomena. The classification of processes offered here is a clear-cut one based on quantitatively observed data, organized by means, sigmas, and correlation coefficients, etc. It is general to all the social sciences, thus having an applicability that is universal to society.

It is submitted in favor of the S-system of hypotheses that the approach to the analysis of the phenomena of society is a realistic one. It deals with data on a level where they can be objectively and reliably dealt with, and with data that are in constant, practical, everyday use. Sociologists have been inclined to try to deal with somewhat nebulous and intangible processes which have attractive emotional and evaluative aspects, such as liberation, estrangement, anticipation, ossification, telesis, decadence, consciousness of kind, institutionalization, simulation. (See Ref. 25, p. 44.) The description of these may prove eventually to be useful hypotheses. But at an early stage of development of a science, measurable processes such as our competing, mobility, rescheduling, and the like may be a more solid foundation to build upon by observing, relating, understanding, and increasingly predicting and controlling them in general and in differential fashion for subforms of differing content. The phenomena measured by rescheduling, for example, are among the most important and widely used of any in life. Every person in planning or implicitly expecting his day's program of activities, every business or other human organization in arranging its activities ahead, is actually scheduling and rescheduling events and durations between them. Dealing with these processes for which data are immediately at hand in measurable form, and which are socially important and widespread (though deficient in emotional and evaluative glamour), is what is meant in this volume by a realistic societal analysis.

### III. COMPLEX PROCESSES—MOMENTUM OF CHANGE, $\mu(I)P$

#### *A. Definition and Examples*

Thus far in exploring the chief equations of change deducible from the basic S-theory equation, only the amount of a change

and the velocity of it have been considered. The societal momentum <sup>60</sup> of any change is to be considered next.

When the third index, the population, is multiplied by the two indices of a velocity, the product is the *momentum* of a social change:

$$(I)T^{-1}P = Mm = \text{momentum of a societal change} \quad (\text{Eq. 78a, Ch. X})$$

This expresses the extensity of a change, the number of people who are changed (P), as well as the intensity or amount of the change (I), and its swiftness ( $T^{-1}$ ). A social movement may increase in momentum by increasing either its clientele or their rate of change. Thus the "diffusion of culture" varies with the culture complex ( $I_i$ ), its rate of diffusion ( $T^{-1}I$ ), and the size of the population (P) through whom it is diffused. As an example, the diffusion of missionary religious culture complexes such as Christianity, Islam, Buddhism, and Bahaism can be studied: *first* considering the system of traits constituting each complex; *second* considering the time taken to convert a people in different centuries; and *third* considering the number of people converted in different parts of the world.

Momenta thus studied state, in a systematic way, effects more than causes. Causal analysis remains to be done to explain the observed momenta. But accurate knowledge of results is essential in studying causes.

Thus the necessary causal steps to be taken in any field, for example education, depend upon knowing the current momentum of the educational process. Its curricula and examinations define the kinds and amounts of change,  $I_i$ , to be achieved by the youth; their scheduling at successive age levels defines the average velocities; and the number of pupils in the nation completes the definition of the educational momenta of all kinds.

$$Mm = \text{t}T^{-1} : IP_i \quad (\text{Eq. 78b, Ch. X})$$

Society's administrators in planning budgets, personnel, and policies or legislation, need to consider all three factors of momentum, separately and jointly. A joint index such as Mm makes comparable such movements as adult literacy campaigns, where literacy up to a defined standard (as a score of 'I in some reading

test) may be produced twice as quickly among the masses of region P as in region Q, but which teaches only one third the population taught in region Q. In which of the two regions is a given outlay of funds more efficiently administered (assuming other factors equal)? The answer is:

$$1.5 \text{ Mm}_P = \text{Mm}_Q \quad \text{for} \quad (\text{Eq. 79a, Ch. X})$$

$$\left(\frac{\text{IP}}{\text{T}}\right)_P = \left(\frac{\text{I} \times 3\text{P}}{1.5 \times 2\text{T}}\right)_Q \quad (\text{Eq. 79b, Ch. X})$$

showing that the literacy momentum of region Q is 50% greater than in region P. Thus momentum, in controlling societal evolution, provides a key concept for planning the national effort needed to achieve some goal such as complete literacy by a scheduled date.

### B. Imitation

Two important varieties of momenta, much discussed in sociological and psychological literature, are "imitation" and "social inheritance." *Imitation* having aspects of amount, speed, and widespreadness comes under the definition of Mm with its intensity, protensity, and extensity factors ((I), T, and P). Imitation is momentum of a change coming over a plurel from another contemporary plurel. The East imitates the European dress. The momentum in this case is measurable through surveys counting the proportion of an eastern native population that wears varying numbers of standardized European items of dress replacing indigenous costumes (such as felt hats, no veils, coat, vest, long pressed trousers, etc.). A "Europeanization of dress" score could be readily worked out for the I and determined on sample populations at two or more dates, yielding the index for this one type of momentum of intercultural interaction.

For further and more rigorous evidence, exploring the use of S-theory to put hypotheses into a more exact form for testing out, Tarde's hypotheses (often miscalled "laws of imitation") may be briefly restated as hypotheses in S-notation.

Let:

<sup>p</sup>P<sub>p</sub> = the population imitated

<sup>p</sup>P<sub>q</sub> = the population imitating

I<sub>s</sub> = the culture complex (collection of I traits with any sub-classification) which is imitated

$^M I_r$  = another culture complex,  $|r$ , reduced to a single mean score on a scale in which superiority and inferiority are imputed (see Ref. 15)

Tarde's first hypothesis (as quoted in Ref. 41, p. 240) is:

1. "The inferior imitate the superior."

Stated in symbols this is:

$$Mm_{s,q} \cdot ? = T^{-1} I_s P_q, \quad \text{where} \quad {}^M(I_p) > {}^M(I_q) \quad (\text{Eq. 80, Ch. X})$$

As usual the question mark before the equality sign denotes an hypothesis, an equation still open to question and calling for an answer as to its truth.

This equation may be expanded in words as: the momentum of change of the culture complex  $I_s$ , in population  $P_q$ , is equal (according to this hypothesis) to the product of the increment in the indicants (of the culture complex derived from the superior group) and the population of the inferior group that is imitating, divided by the time in which the imitation occurred. More briefly, the momentum is the velocity of imitating the superior group times the number of inferior people imitating. The  $Mm$ -units are indicant-persons-per-time units.

For verification of these hypotheses more studies are needed determining in particular situations the  $T$ ,  $I$ , and  $P$  factors, devising scales of inferiority-superiority, determining the amounts of inequality needed to produce given velocities, or affect given numbers of people. A simple type of study is to take one attribute ( $I^0$ ), such as an unusual garment worn by the Prince of Wales on some foreign tour, and determine sales (roughly =  $P$ ) of it in a period ( $T$ ) thereafter.

A measure of prestige is the ratio:

$$P_{q/p} = \text{a "prestige" ratio} \quad (\text{Eq. 81, Ch. X})$$

as this is greatest when the imitated plurels,  $P_p$ , reach their lower limit of one plurel and one person, and  $P_q$ , the imitators, are numerous. This prestige ratio shrinks as the imitated group in the denominator becomes larger and the imitators in the numerator become fewer. Compare the prestige of a Gandhi or a Hitler with the prestige of ordinary Europeans in some tropical area,

each measured by their ratio to the number of their followers (or imitators in defined respects).

Tarde's second hypothesis is:

2. "In the absence of interference imitations tend to spread in a geometrical progression."

$${}^i\text{Mm?} = {}^a\text{Mm}I^{i-1} \quad (\text{Eq. 82, Ch. X})$$

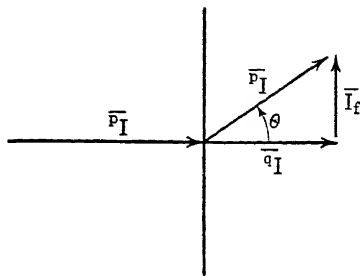
where  $i$  denotes the  $i$ th term in the series starting from an initial momentum,  ${}^a\text{Mm}$ , and  $I$  is the common ratio.

The progression may be due to temporal acceleration, or to a geometric ratio in the successive numbers of imitators, or to both. The nature of "interference" for diverse types of imitation needs study. The equation squarely presents the research problem of determining the numerical size of this ratio in specific situations and noting trends or norms from which to predict the future.

Tarde's third hypothesis is:

3. "Imitations are refracted by their media."

In the accompanying figure let  $\overline{PI}$  be the vector representing the characteristic among the persons imitated, and let  ${}^a\overline{I}$  represent it among the persons imitating,  ${}^aP$ . The angle  $\theta$  is the angle of refraction, as when the English language becomes the American version in the United States, or "pig-



din" English in Chinese ports (when it contacts and is deflected by the local factors,  $\overline{I}_f$ ). As usual in direction-cosine theory, which is the geometric representation of the S-theory of this volume, the  $\theta$  is specified by its cosine which is a function of the ratio  ${}^aI/\overline{P}$  in the figure, and is also the correlation coefficient,  $r$ , between the original and the refracted culture complex. As the two variables in this  $r$  exist in different groups, the  $r$  cannot be calculated by the usual methods. But recalling that  $r$  is interpretable as the geometric mean of the two percentages of common

elements in the two variables, suggests using this index to measure the refraction.

$$\sqrt{({}^cI/{}^pI)({}^cI/{}^qI)} = \cos \theta \quad (\text{Eq. 83, Ch. X})$$

Thus the part of the imitated culture complex or traits that is exactly reproduced is reduced to an indicant,  ${}^cI$ . The whole of the culture complex imitated is reduced to an indicant also,  ${}^pI$ , as is also the whole culture complex of the imitators,  ${}^qI$ . Technics for such quantifying of culture in a single number have been developed. (See such texts as Lundberg, G. A., *Social Research*, Longmans Green, for an exposition of principles. For specific examples see Chapin's "Socio-Economic Status Scale," Dodd's "Village Hygiene Scale," Leahy's "Urban Home Environment Scale," the Rural and Urban Public Health Appraisal forms, and other scales referred to in Chapter III.) The ratio of those indicants is a proper fraction, as the common complex is a part of each whole complex.

The square root of this fraction is the cosine  $\theta$ , giving the angle of refraction. Note that this equation is not an hypothesis. It is a definition. It does not state what the refraction should be found to be under certain circumstances. It states an index for measuring the amount of refraction even if it is zero, as when a culture complex is so well defined as to be copied perfectly, such as the international playing of contract bridge, chess, or tennis.

The indicants connected with refraction may be tabulated thus:

REFRACTION INDICANTS

Per cent of elements imitated	$\cos \theta = r$	$\theta$	Meaning	Imitated plurel	Imitating plurel
100	1.00	0°	No refraction		
50	.71	45°	Half refraction		
0	0	90°	{ Full refraction } { Zero reflection }		
*	-.71	135°	Half reflection		
*	-1.00	180°	Full reflection		

Full refraction would denote that the imitated culture complex had been so completely and unrecognizably changed as to show no correlation in the imitated and imitating plurels. Full reflection would denote that the culture complex had been completely reversed, i.e., its opposite had been adopted, as when a Greek

conqueror of Jerusalem ordered swine to be sacrificed on the altar of the temple of the Jews to whom swine were defiling and tabooed.

Bogardus adds a fourth hypothesis as follows:

4. "Conscious imitation tends to vary with imputed superiority."

$$Mm_q \bullet (I')_{p-q} = r? > 0 \quad (\text{Eq. 84, Ch. X})$$

The hypothesis is that in the plurels  $P_q$  the momentum of conscious imitation correlates with the difference on some scale ( $I'$ ) of imputed superiority-inferiority between the position of the imitated ( $|_p$ ) and the position of the imitators ( $|_q$ ). Quantitative studies of the degree of correlation existing here under various conditions are lacking. The equation provides an excellent subject for research. How high, under specified conditions, is the correlation,  $r$ , asserted by the phrase "tends to vary"?

A fifth hypothesis, suggested by Bogardus, is:

5. "Conscious imitation correlates with social nearness of the stimulus."

$$Mm \bullet I_D = r? < 0 \quad (\text{Eq. 85, Ch. X})$$

The hypothesis here is that the momentum in any plurel of conscious imitation correlates negatively with the social distance test score,  $I_D$ , of that plurel from the group or stimulus imitated.

6. A further hypothesis in the field of imitation is that informational culture traits tend to have greater momentum, their corresponding behavioral traits tend to have less momentum, and their corresponding material traits tend to have least momentum.

$$Mm' = Mm''J = Mm'''JK \quad (\text{Eq. 86a, Ch. X})$$

where

$Mm'$  = momentum of an idea, or knowledge

$Mm''$  = momentum of its corresponding habit, or behavior

$Mm'''$  = momentum of its corresponding equipment, or material culture traits

$J$  = a fractional coefficient between 0 and +1, i.e.,  
 $1 > k > 0$

$K$  = another coefficient between 0 and +1, so that

$$Mm' > Mm'' > Mm''' \quad (\text{Eq. 86b, Ch. X})$$

This means that ideas travel furthest and fastest, and then habits change with a lag, and finally copies of material things are reproduced last in a given culture complex. Knowledge about guns, autos, or coeducation reaches a remote and isolated people before they are used, and their use in imported form tends to precede their local production.<sup>61</sup>

Care must be taken in testing this hypothesis not to compare information in one complex with behavior in another, or equipment in a third. Each trait and complex has its own and differential resistances, so that the velocity of penetration of the use of a gun may greatly exceed the velocity of the idea of coeducation. This apparently denies the hypothesis, as an idea's momentum is here less than a habit's momentum, but the phrase "their corresponding" has been neglected. To be sure trader's goods may penetrate further than the trader or his ideas, but the hypothesis merely affirms that knowledge of the function of *those* goods precedes *their* adoption, and this precedes *their* local production. The importation of the physical goods accelerates the diffusion of knowledge about them that must precede their use. An implication of this hypothesis, if true, is the power of radio broadcasting and television in preparing the way for any change of habits or material culture which the controllers of a society may wish to achieve. Centralized governments such as Russia and Germany and others today (1939) are using this principle effectively for molding the total national culture along lines desired by the Government.

These hypotheses on imitation are perhaps enough to illustrate a large field for further research. Hypotheses need to be invented, precisely formulated by the aid of such tools as this S-theory, and tested against data to discover the degrees to which, and the conditions under which, the hypothesized generalizations are valid.

### C. Social Inheritance

A second important variety of momentum is *social inheritance*. This is the changing, or educating, of each generation of babies to the culture of the parent generation of their in-group. Here the time,  $T$ , involved in the formula  $Mm = T^{-1}IP$ , is a generation, and the indicant is all  $i$  culture traits of a given culture com-

municated in varying degrees of completeness from the former ancestor-worshiping-China to rapidly changing immigrant sections of New York. The P is all the births or immigrants coming into a cultural plurel. Obviously the momenta of social inheritance vary greatly with the varying of its three factors.<sup>62</sup>

These principles are used in social control, though often not as explicitly as is possible by a knowledge of the formula for Mm.<sup>63</sup> Thus at the end of the World War the United States realized that the momentum of assimilating its foreign-born immigrants was too little and acted to increase it by: (a) Americanization schools and programs to increase the velocity of assimilation; and (b) restricting immigration by quota laws to reduce the annual new population to be assimilated. Thus existing agencies were able to reach an increased proportion of the admitted foreign-born.

$$+Mm = IT^{-1}P_q, \quad (\text{Eq. 88, Ch. X})$$

$P_q$  = recent immigrants (imitators)

T = period from admission to assimilation

I = indicator of assimilation, as in completion of naturalization, or as in passing an English reading test.

#### *D. Summary of the Single Processes, Both Simple and Complex*

Herewith the exploration (for the limits of this volume) of the single processes has been completed. Before describing the aggregated processes, the classification of all S-theory processes may be clarified by tabulating them in the following table. Following the outline of this chapter the single and the aggregated processes are first differentiated in columns. Then within the single processes, the classes are the four sectors I, P, T, L in four columns. Within each column the rows represent the three commoner exponents classifying processes as nullary, primary, or secondary. Within each resulting cell the process with some chief sub-processes (in parentheses) are given by name and by Brief-S formula. The main processes, being highly generalized forms of change, have names which for the most part are new terms in Sociology. But when these forms are combined with more specific content, such as desiderata, the processes described become such familiar ones as "co-operating," "accommodating," "conflicting," "competing," and others.



$ ^1 = 2$ <i>Secondary processes</i> Variances, $\Sigma(\Pi')N^{-1}$ Second-order moments	<i>Relations</i>  (some subtypes) $\rightarrow$ change in the distribution of: $\rightarrow$	Reindicating = ${}_tI^{+2}$  (Competing = $Cp^2 = .5\Sigma_i V^2$ characteristics)	Repopulating = ${}_tP^{+2}$  (Mobility = $Mb^2 = .5\Sigma_i P^2$ population)	Redurating = ${}_tT^{+2}$  (Rescheduling = $Rs^2 = .5\Sigma_i T^2$ durations)	Resapiating = ${}_t(L)^{-2}$  (Remapping = $Rm^2 = .5\Sigma_i (L^2)^2$ spaces)	Varying Opposing $ _1 = 2$ " " " " " " Exploiting $ _1 = 3$ Sporting " Evolving $ _1 = 5$
$ ^1 = 3$ <i>Tertiary processes</i> Skewnesses, $\Sigma(\mathcal{I}')N^{-1}$ Third-order moments	Symmetric relations	Skewing = ${}_tI^3$	P-Skewing = ${}_tP^3$	T-Skewing = ${}_tT^3$	L-Skewing = ${}_t(L)^3$	
Additional <i>Complex processes</i> * Indices, (I)	Sums, differences, products, ratios, of single simple processes above	Momentum = $Mm = (PIT^{-1})$  Tensing = ${}_tE = (PDV^{-1}T^{-1})$	Complexes of ${}_tL^1$ , lines, (points, lines, areas, volumes) and of P, I, and T such as Densing = ${}_t(PI)^{-2}$			

\* "Simple" is the adjective contrasting with "complex", as "single" contrasts with "aggregated". Complex processes denoted by indices (I) are additive (including multiplicative) combinations of simple processes. Mean primary processes, secondary, and tertiary processes involve indices in a ratio of 1 to 1' and are, therefore, complex processes.

The process formulae are here shown for any series of periods,  $t$ , in more general form than the process for one particular period,  $t'$ .

IV. AGGREGATED PROCESSES,  $\iota(I)_i$ 

## A. Definition of Aggregated Processes

The preceding single processes, definable by a single index, have been described and illustrated by graphs and other quantified situations quoted from the current literature of the social sciences. Some aggregated processes will be outlined now as suggestive hypotheses for further research. No recorded quantitative situations have been discovered as yet by the author for illustrating most of these aggregated processes.

An aggregated process is one which by the rules of S-notation requires more than one index to define it. The manner of combination of the indices in the compound is not clear as yet. It may be an aggregative combination as in a set of indices compared for a common series of periods in one population, as in S. 23, Ch. XII showing the social trends in the Soviet Union. It may later prove to be expressible as a summative combination, as in combining indices into some form of single weighted arithmetic average.<sup>64</sup> Consequently the semicolon symbol is used to denote any unspecified one of the various mathematical operations of combining (+ -  $\times$  /  $\bullet$  : ::).

Aggregation is the loosest, most general, and least specific form of combination, so that it will be assumed here, unless contrary indication is given.

$\iota(I)_i$  = an aggregated process, change in an aggregation of indices  
( $|_i$  in number, in periods  $\iota|$ ) (Eq. 89a, Ch. X)

$\iota(I)_{\Sigma i}$  = a summative process, change in a sum of indices, a polynomial (Eq. 89b, Ch. X)

## B. Illustrative Hypotheses towards Measuring Aggregated Processes

## 1. UNIFYING AND SEPARATING

When a population composed of plurels with non-overlapping membership is breaking up into a larger number of plurels, dissociating is taking place. If simultaneously the separate plurels are developing different characteristics (as is natural whenever there is isolation or antipathy between the plurels), then dissimilarizing is also going on. Sometimes it may be difficult to distinguish these two processes, and a term for their joint occur-

rence may be useful. Thus, if more fraternal lodges are being formed in a community with increasing diversity of their characteristics, this joint process (which may be termed "unifying-separating") is going on, for if every person belongs to one and only one lodge, then new lodges can only be formed (without newcomers from outside of that population) by taking members away from existing lodges—which is clear-cut dissociating. On the other hand if every person is a member of a lodge, and new lodges are formed by multiple membership of persons in more than one lodge, then differentiating is going on. But usually both processes are mixed (often together with some "recruiting" from newcomers).

$$(+_tP_{2p}^0); \quad (+_tI_{2i}^0)? \quad = \quad + (I)_{sp} \quad (\text{Eq. 90a, Ch. X})$$

Dissociating and differentiating = separating

When the double process goes in the opposite direction with groups merging and becoming culturally assimilated, the combination of associating and assimilating might be termed "unifying."

$$(-_tP_{2p}^0); \quad (-_tI_{2i}^0)? \quad = \quad - (I)_{sp} \quad |^s = 9;0;0;0$$

Associating and similarizing = unifying (Eq. 90b, Ch. X)

Whether such a concept will prove useful, or be forgotten, depends on whether the two constituent processes can be measured in one situation.

## 2. ATTRACTING AND DETRACTING

When a group is recruiting, enlarging its membership, and simultaneously the members are intensifying their desires for whatever the group represents to them, the two processes of recruiting and evaluating are going on. Their combination which might be termed "attracting" is a double process commonly characterizing the growth of any group. Its opposite would be conflicting and devaluating in a process of "detracting."

$$(\pm_tP_v); (\pm_tI_D) \quad ? \quad = \quad \pm (I)_{At} \quad |^s = 9;1;0;1$$

Recruiting and evaluating = attracting  
 Conflicting and devaluating = detracting (Eq. 91, Ch. X)

During a conflict the double process becomes mixed, in that each group is recruiting and evaluating (intensifying its loyalty)

within itself while attempting to eliminate population and to devalue in the opposing group.

After the conflict detracting is more apt to characterize the group losing out in the conflict, for then the number and the loyalty of its membership tend to decrease, while for the victor attracting tends to go on again.

In weighing these hypotheses, the reader must continually be on guard not to be influenced by the connotations of the *names* of the processes, but to think whether what the *formulae* symbolize is a true or useful description of phenomena. The names are aids to reflection for those not facile in thinking in terms of these unfamiliar S-symbols, but they may also prove ensnaring unless the reader is on guard. Thus, the terms "expanding" and "contracting" could be used for  $\pm(I)_{At}$ , but with differing connotations. The essential question is: Does  $+(I)_{At}$  describe the increasing of population and of ambitions (as in an expanding imperialist country), and does  $-(I)_{At}$  describe the reverse (as in war casualties and waning of imperialist ambitions when an inconclusive war drags on)? What are the conditions under which this process exists, accelerates, decelerates, or reverses in direction? Are the conditions generalizable, or is the symbolizing of the process in a single index deceptive in that the process and the conditions for it are numerous and specific to each type of situation?

### 3. COERCING

A large number of societal processes may become measurable when analyzed into terms of positive and negative desiderata, V, intensities of desire, D, plus probabilities and time sequences. Concepts such as coercion, liberation, persuasion, threatening, punishment, repression, and others can be so reduced. For one example, consider coercion. Two instances will make the abstract symbols more vivid. The Spanish Inquisition coerced the heretic with the alternative:

"Persist in your belief ( $+{}^tD_{,,}$ ) and be tortured to death ( $-{}^tD_{,,}$ )  
or  
"Recant ( $-{}^tD_{,,}$ ) and be set free ( $+{}^tD_{,,}$ )."

The coal company with its company store coerces the miners with the choice:

"Buy cheaper elsewhere ( $+{}^tD_{,}$ ) and lose your job ( $-{}^tD_{,}$ )"

or

"Buy expensively here ( $-{}^tD_{,,}$ ) and keep your job ( $+{}^{t''}D_{,,}$ )."

The essential pattern here can be described by the matrix

$$\begin{bmatrix} +{}^tD_{,} & -{}^{t''}D_{,} \\ -{}^tD_{,,} & +{}^{t''}D_{,,} \end{bmatrix} = {}^tD_i \text{ (Eq. 91.5a, Ch. X)}$$

where there are:

1. alternative acts ( $|$ , or  $|_{,,}$ ) positively and negatively desired ( $+D$ ,  $-D$ ) by the party coerced;
2. each act having a high probability, imposed by a second party, of a consequent event  ${}^t|$ , of opposite valuation ( $-$  or  $+$ ) to the first party;
3. the relative valuations,  $D$ , of the four desiderata being of such sizes that the first party chooses "the lesser of two evils" in acting against its desire in order to avoid a worse consequence.

Equations of the ratios formed from the four elements of this matrix can be built up and manipulated in various ways. Thus if the aversion ( $-D$ ) to death is equal to the aversion to recanting, their ratio equals unity. This is the critical point of indecision or balanced choice; if either aversion changes, the equation may become an inequality and measure the compliance or defiance of the coerced party.

$$-{}^{t''}D_{,}/-{}^tD_{,,} = 1 = \text{a measure of the effectiveness of coercing} \\ \text{(Eq. 91.5b, Ch. X)}$$

The coercing may vary in three ways:

- a. according to the absolute size of the smaller negative desire  $| -{}^tD |$  the coercion varies in *importance* from trivial to grievous;
- b. according to the relative size of the two alternative negative desires,  $-{}^{t''}D_{,}/-{}^tD_{,,}$ , the coercing varies in *effectiveness* from completely effective to ineffective, as when this ratio drops below the critical point of unity and the first party defies the second, preferring death to recanting, or being fired to being exploited, and is no longer "coerced";

- c. according to the probability of the second party imposing the indicated consequence on the first, the coercing varies in *probability* from an idle threat to an inescapable consequence.

Recalling that joint probability is but the simpler form of correlation of all-or-none indicants, the result of this analysis is that the three aspects of importance, effectiveness, and probability of the coercing process may be measurable by an amount, a ratio, and a correlation, respectively, among the intensities of desire,  $D$ , of the coerced party in the situation.

For suggestions to be explored, "liberating" may prove to be definable and measurable as decreasing towards zero the probability of the undesired consequence in a situation of coercion; "threatening" as coercing with lower or unknown probability; "persuading" as prior acts by the second party which correlate with subsequent valuating (changing intensities of desires) in the first party; etc.

#### 4. TOLERATING, NON-TOLERATING, AND PROSELYTIZING

In toleration the basic situation is that two parties are in conflict over some desideratum, and at least one tends to eliminate the other. A church may want all Christians to believe the orthodox creed to the point of burning at the stake those heretics who refuse to believe. But the essence of toleration is that in such a situation the parties reduce their desires for universal possession of the desideratum and their supplementary desire to kill off the opponent. They are basically in conflict, yet they accommodate to each other; they live and let live in spite of antagonism. Although the formula is the same as for "detracting" above, the time relationships of the elements differ. In "detracting," full conflict *has been going on* and desire wanes; whereas in "tolerating" desires are reduced while the conflict is incipient, i.e., before it has reached the stage of actual killing or eliminating of opponents from the field of that desideratum.

The opposite of tolerating is where the zeal for universalizing the desideratum increases, and recruiting of membership into the plurel identified by that desideratum is going on. In the field of religious beliefs this is termed proselytizing, the intensifying of

desire to spread the beliefs (desiderata) of one's religion and the winning of converts. But "proselytizing" in its root meaning of "coming towards," a new belief or valuation, might serve as a generalized extreme deviation from tolerating. Intolerance, then, would denote the relationships in the process where conflicting was going on unchecked, and desires (for the value at stake in the conflict) were becoming more intense. Tolerating is intermediate between proselytizing and intolerating, since in tolerating, conflict (the elimination of the opposing population) is held to a zero amount, in spite of an intense desire to universalize the desideratum at stake.

$$\begin{aligned}
 (+_tP_v); (+_tI_D) ? &= +(I)_{T_1} = \text{"proselytizing"} \quad (\text{Eq. 92a, Ch. X}) \\
 &= \text{recruiting and desiring} \\
 ({}_tP_v = 0); (+_tI_D) ? &= (I)_{T_1} = \text{"tolerating"} \quad (\text{Eq. 92b, Ch. X}) \\
 &= \text{desiring without recruiting or eliminating} \\
 (-_tP_v); (+_tI_D) ? &= -(I)_{T_1} = \text{"intolerating"} \quad (\text{Eq. 92c, Ch. X}) \\
 &= \text{desiring and eliminating opponents}
 \end{aligned}$$

## 5. SOCIALIZING VS. INDIVIDUALIZING

Another hypothesis towards reducing a compound process to measurement concerns socializing. The essence of socializing is, in Ross' words, "the development of the we-feeling in associates and their growth in capacity and will to act together." (Ref. 60, p. 375.) In Von Wiese's words it is the establishing of "ethically sanctioned intragroup and intergroup bonds." (Ref. 78, p. 367.) In the system of concepts developed in this volume, if the desideratum,  $V$ , represents the "we-feeling," the "group bonds," then its increase is one form of co-operating,  $+_tV$ . At the same time all mutually incompatible desires, as for one's own way of doing things, must be accommodated,  $-_tD_v$ , if there is to be the "capacity to act together."

Thus co-operating and accommodating together, when each is limited to specific types of desiderata as suggested above, may constitute the meaning of socializing.

The paired opposite would be a decrease of the we-feeling, the attitudinal desideratum binding the group together (measurable by some attitude test or indicators of cohesive behavior), accompanied by increasing desire to have one's own way and not ac-

commodate to others. This may be dubbed "individualizing," pending more rigorous analysis and trial in measured data. This process is a specific form of the regressing and evaluating processes.

$$(+_tV_{Co}); (-_tD_s) ? = + (I)_{sc} = \text{"socializing"} \quad (\text{Eq. 93a, Ch. X})$$

= specified co-operating and accom-  
modating

$$(-_tV_{Co}); (+_tD_s) ? = - (I)_{sc} = \text{"individualizing"} \quad (\text{Eq. 93b, Ch. X})$$

= specified destroying of group de-  
siderata and intensifying of indi-  
vidual desires

When the specification of the desiderata is altered, the process described in the sociological literature under the rubric of "adjusting" seems to be described by these symbols. In adjusting the parties accommodate to each other in reducing any desires that are unsatisfiable in that situation, as well as in acting so as to increase the desiderata in common that are possible of gratification. A new boy at boarding school, the delicately reared pioneer's wife in the wilderness, the acquiescent people of a conquered territory, all adjust to their new situation by reducing whatever desires are physically and socially impossible, and by doing whatever they can to get along with their new associates and build up whatever desiderata they may have in common. (This analysis, though inadequate, may be suggestive.) The formula for adjusting is the same as for socializing, but the definition of  $D$  differs, as in "adjusting"  $D$  also includes desires with reference to the physical environment, whereas in "socializing"  $D$  denotes desires of a social sort, i.e., desires for values of interaction with other people. Also in "socializing," the desideratum,  $V_{Co}$ , is group solidarity, the we-feeling, while in adjusting it may be a more general set of values, desired by all the group in common. The reverse direction of this process might be labeled tentatively as "maladjusting." This paired process might be symbolized by:

$$\pm (I)_{Aj} = \text{adjusting and maladjusting} \quad (\text{Eq. 93c, Ch. X})$$

The form of the formula is that of Eq. 93, Ch. X above, but the class scripts on  $V$  and  $D$  would be altered to denote a different set of desiderata.

In a similar way the same formula might be refined so as to symbolize the process of "adapting." Here the desiderata are other-than-social. Thus "adjusting," if these definitional hypotheses are confirmed, would be subdivided into "socializing" with reference to interhuman desires, and "adapting" with reference to other desires.

## 6. VARIATING AND SIMILIZING

In Chapter V "variation" was defined as the combination of differentiation and dispersion. The process of varying may then be defined as the combination of dissimilarizing (differentiating) into more kinds of characteristics and dispersing into larger standard deviations of each characteristic. Varying is thus increasing the number of qualities and the range of the quantity of each quality. Biological and sociological variation includes both types of differences. North's social differentia of rank, function, culture, and interest (Ref. 50) are both qualitative and quantitative and would, therefore, be examples in their dynamic aspects of our compound process of varying and its oppositely directed process, which might be termed deviating or "similizing."

$$+_t(|z_i); +_t(T) ? = (I)_{vr} = \text{"varying"} \quad (\text{Eq. 94a, Ch. X})$$

= dissimilarizing qualities and dispersing quantities

$$-_t(|z_i); -_t(T) ? = -(I)_{vr} = \text{"similizing"} \quad (\text{Eq. 94b, Ch. X})$$

= qualitative similarizing and quantitative equalizing

The manner of combining the two constituent processes is a problem for research. Can suitable units and weighting be devised to make adding them yield an index reflecting changing of either the qualitative or the quantitative type, or should the combination remain an aggregation requiring more than one index to express, as in the aggregative formula:

$$(I)_{vr} ? = {}_t(T_i) \quad (\text{Eq. 94c, Ch. X})$$

or is some other mathematical function more useful?

## 7. OPPOSING

Often the processes of competing and conflicting coexist and are difficult to distinguish. This joint process might be termed

"opposing." Since competing is a process without an opposite (as a negative standard deviation is debarred <sup>65</sup>), opposing has no fully reversed process in which all its constituent processes go in the other direction.

Opposing is symbolized by a primary and secondary process as:

$$(-, P_v); C_p ? = (I)_{op} = \text{"opposing"} \quad (\text{Eq. 95, Ch. X}) \\ = \text{conflicting and competing}$$

To measure this, as well as any other compound process, the constituent processes must be measured, of course.

## 8. EXPLOITING

As an example of an hypothesis of a trinomial process the following tentative analysis of the process of exploiting is offered.

It is suggested that "exploiting" involves coercing and competing in a situation of dispersion. The exploiting party is competing with the exploited parties; there is a large gap in their economic power, i.e., the sigma of this characteristic is large; and the exploiting party coercively links lesser negative with larger positive values for the exploited party. The exploiting plantation owner in competing with his native employees for the larger share of the plantation's proceeds from sales has a large gap in economic power separating him from them, and so is able to link inseparably the negative value "inhuman working conditions" with the greater positive value "livelihood at all."

Symbolically the aggregation of the three constituent processes is:

$$Cr ; C_p ; (^\circ I) ? = (I)_{Epl} = \text{"exploiting"} \quad (\text{Eq. 96, Ch. X}) \\ = \text{coercing, competing, and dispersion} \\ \text{larger than average}$$

An hypothesis such as this may be verified by multiple correlation technics. Suppose that in each of a hundred situations, involving exploitation to varying degrees as rated by a group of sociologists trained in this field, the amount of coercing, competing, and dispersion were separately measured, each as defined above. Then the multiple correlation coefficient between the three constituent processes and the ratings of exploiting is calculated. Proportionately as this multiple correlation, corrected for attenuation, approaches unity, exploiting is proved to be com-

posed of these and only these processes. Eq. 96, Ch. X can be made into the multiple regression equation by giving each constituent process its multiple regression weighting, and by including a constant term on the left to adjust for the origins chosen. If, as is probable with crude technics at first, the multiple correlation is significantly below unity, two methodological choices present themselves. The first choice is to say that residual unmeasured processes are constituents of exploiting (to the extent of the multiple coefficient of non-determination,  $k_{0 \cdot 123 \dots n}^2$ ,  $(= 1 - r_{0 \cdot 123 \dots n}^2)$ ) over and above the processes of coercing, competing, and the dispersion. The research problem is then to discover and measure these residuals until the multiple correlation approaches unity (within limits of error) and the validity of our analysis of "exploiting," as measured, is proved. The second choice is to say that Eq. 96, Ch. X defines a process, denoted by the symbol on the right, which has been inappropriately labeled "exploiting." A new term might be coined and the fact recognized that this new process correlates to the extent  $r$  with the process of "exploiting" (as measured by the ratings based on a schedule card objectifying conventional notions of exploitation.<sup>66</sup>

## 9. EVOLVING

A final hypothesis of an aggregate process, analyzed into what may or may not prove to be its necessary and sufficient constituent processes, is the quinnomial hypothesis towards defining the societal process of evolution. Building on the biologists' analysis of evolution into the processes of variation, selection, and heredity, and expressing these in standard S-notation gives:

$$Vr; Op; Sb ? = (I)_{Evol} = \text{"evolving"} \quad (\text{Eq. 97a, Ch. X})^{67}$$

Here varying is defined above (Eq. 94, Ch. X). Selection is competing for the limited desideratum, "the wherewithal to survive," followed by depopulating the less successful competitors. This is a form of our "opposing" process (Eq. 95, Ch. X) as compounded of competing and conflicting (for specified desiderata in specified populations of course). Heredity is objectively measured as stability, i.e., the revarying of characteristics between generations, that trait being most purely hereditary which changes least from one generation to another. This definition

covers cultural as well as germ-plasm inheritance. Therefore, the three processes above break down into five single processes (symbolized by the five-letter subscript "Evolv"):

$$+_{\mathbb{t}}(|_{\mathbb{z}}); +_{\mathbb{t}}(\mathbb{T}); +\text{Cp}; -_{\mathbb{t}}(\text{P}_{\mathbb{v}}), \sigma(\mathbb{t})? = \begin{array}{l} \text{"dissimilarizing"} \quad \text{and} \\ \text{"dispersing"} \quad \text{and} \quad \text{"com-} \\ \text{peting"} \quad \text{and} \quad \text{"eliminat-} \\ \text{ing"} \quad \text{and} \quad \text{"revarying"}? \\ = \text{"evolving"} \end{array}$$

(Eq. 97b, Ch. X)

Such a formula may help conceptually to understand the total evolving of human society. But in any measured instance of course, only a portion of evolving phenomena will be isolated for study—the particular portion being defined by the data from which the indices of the five constituent processes were calculated.<sup>68</sup> \*

#### V. S-SITUATIONS

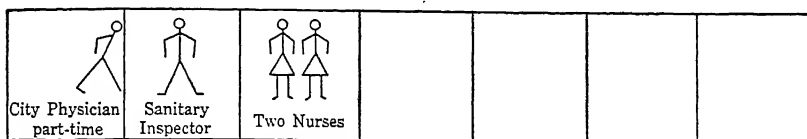
For illustrations of the societal processes which are defined by S-formulae containing  $\mathbb{T}^{-1}$ , the following eighty-five S-situations may be studied. These eighty-five sets of data are arranged to parallel the topics of the chapter. For every situation a process formula is given in the text, but not every formula for a process given in the text has an S-situation illustrating it here. Since the process formulae are hypotheses proposing new operational definitions of concepts, no published data could be found to provide evidence as to the possible utility of many of these hypotheses. They await testing with further research.

As usual each S-situation is accompanied by its descriptive formula which states a particular variation of the fundamental S-theory formula  $S = {}^{\mathbb{s}}(\mathbb{T}; \mathbb{I}; \mathbb{L}; \mathbb{P})^{\mathbb{s}}$  which summarizes all the formulae and quantified data presented in this volume. The quantic number, comprising the exponents of the descriptive formula, follows and classifies each S-situation into one class of societal data as diagrammed in S. 33, Ch. II. The "legend," next, fulfills two functions: (1) to specify the subclass of phenomena symbolized by the S-symbols in a dozen major classes (e.g., in S. 1, Ch. X to specify that the number of dates  $\mathbb{t}$  is "two," the length

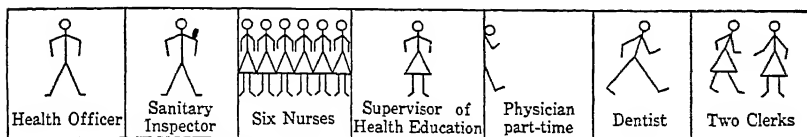
\* For Eqs. 98–102a, Ch. X, see notes at end of chapter.

of each period T is "five years," the particular kind of plurels are "health professionals," etc.) (2) to verbalize the meaning of the S-symbols for the student unversed in the symbols of the S-theory so as to describe the structure of the whole situation as analyzed by the S-formula.

## S. 1



PUBLIC HEALTH STAFF PAID FROM LOCAL FUNDS-END OF 1922



PUBLIC HEALTH STAFF PAID FROM LOCAL FUNDS - END OF 1927

Ref.: Sydenstricker, Edgar, "Safety Always," *Survey*, Vol. LXI, No. 2, October 15, 1928, p. 80.

Descriptive formula:  $S_1 = {}^tT^{-1} : P_p$

Quantic number = 9;0;0;1

Legend:

$S_1$  = The situation  
records for

$T^{-1}$  = a 5-year interval

$P$  = the number of persons  
in each of

${}^t|$  = 2 dates (1922 and 1927)  
with

$|_p$  = 7 professional health plurels

Comment on notation:

The processes of populating ( ${}_tP$ ) and similarizing ( ${}_i|_i$ ) are illustrated in this simple situation:

${}^tP$  = 3.8 persons in 1922

${}^tP$  = 12.5 " " 1927

${}^{t-2}P$  = 8.7 gain in personnel  
=  $8.7/3.8 = 229\%$

$T$  = 5 years

$229/5 = 46\%$  annual adpopulating

Dissimilarizing is measurable by the percentage increase of different kinds of functionaries

$P_p = I^0P_0^{+1}$

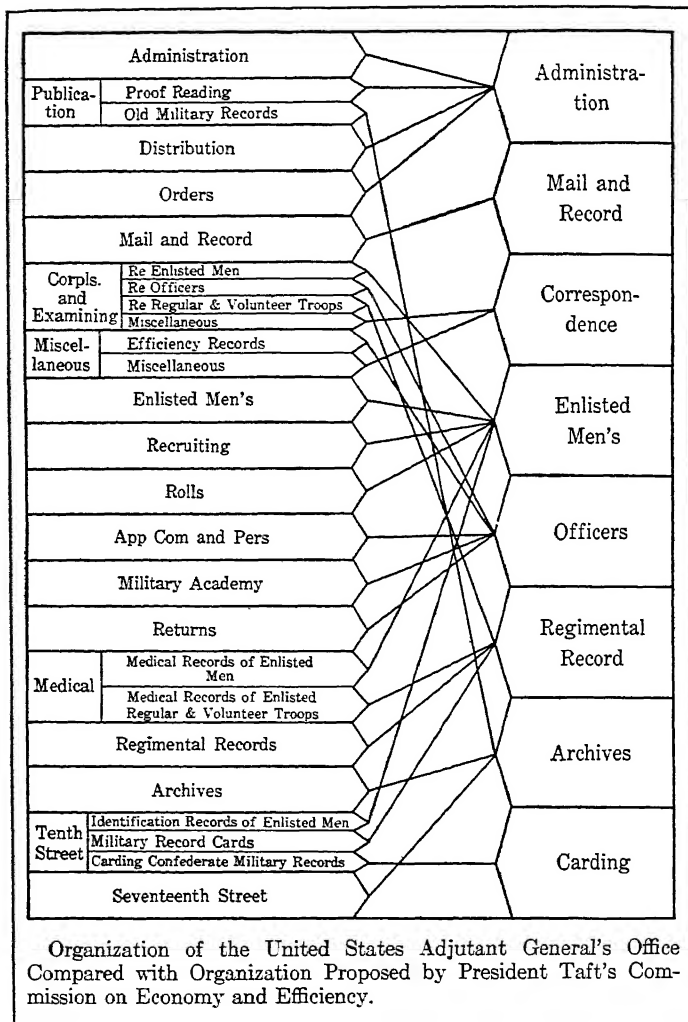
$|_{21} = 3$  in 1922

= 7 in 1927

${}^{t-2}T^{-1}|_{21} = \frac{133\% \text{ increase}}{5 \text{ years}} = 27\% \text{ annual}$

dissimilarizing in relative terms; or in absolute terms, 4 new specialities in 5 years

## S. 2



S. 2 (*Continued*)

*Descriptive formula:*  $S_2 = {}^tT^{-1} : \underline{P}_p$

*Quantic number* = 9;0;0;1

*Legend:*

$S_2$  = The situation

records for each of

$\underline{P}_p$  = the departments of a Government office

${}^tT^{-1} = 2$  dates  $\begin{cases} \text{the present} \\ \text{and after reorganization} \end{cases}$

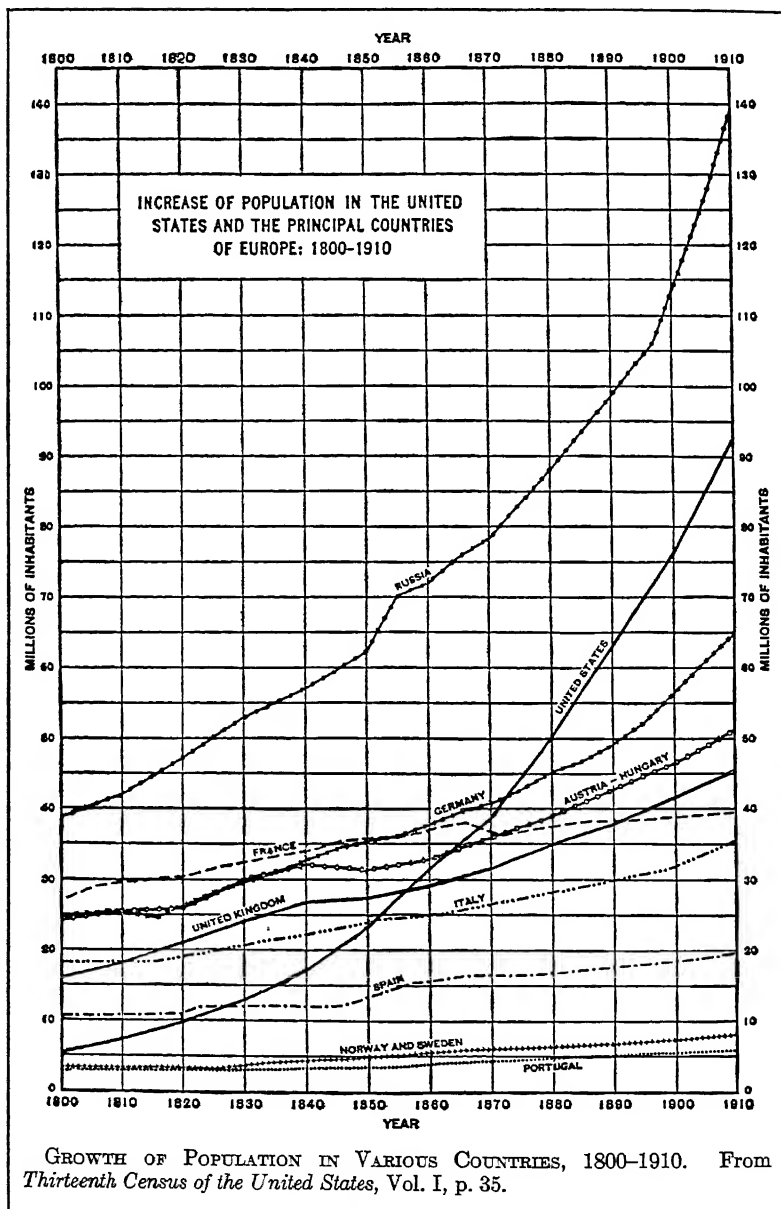
*Comment:*

The process of associating is illustrated on counting the list of departments when these are the population units. The zero exponent reduces the persons in a plurel to a single unit, but on taking more than one such unit the exponent on the population index becomes positive again. The fact that the plurel is the unit, regardless of the number of persons in each plurel, may be written in any of the three equivalent formulae:

Associating =  $-_t\underline{P}_{\Sigma p} = -_t(P^0)_{\Sigma p} = -_t|\Sigma p \quad |^s = 9;0;0;1$  (Eq. 103, Ch. X)

The summation sign is an operational symbol commanding the operator to add up the plurels and thus convert the aggregation in the graph into a single number as in Eq. 103. The 26 departments are merged into 8 new departments. This represents 18/25 of complete association into only one department, or 72% of maximum associating.

## S. 3



S. 3 (*Continued*)

*Descriptive formula:*  $S_3 = {}_tT^{-1} : P_p$

*Quantic number* = 9;0;0;1

*Legend:*

$S_3$  = The situation  
records for each of

'| = beginning in 1800

P = the population

${}_t|$  = 12 periods

of each of

$T^{-1}$  = of a decade each

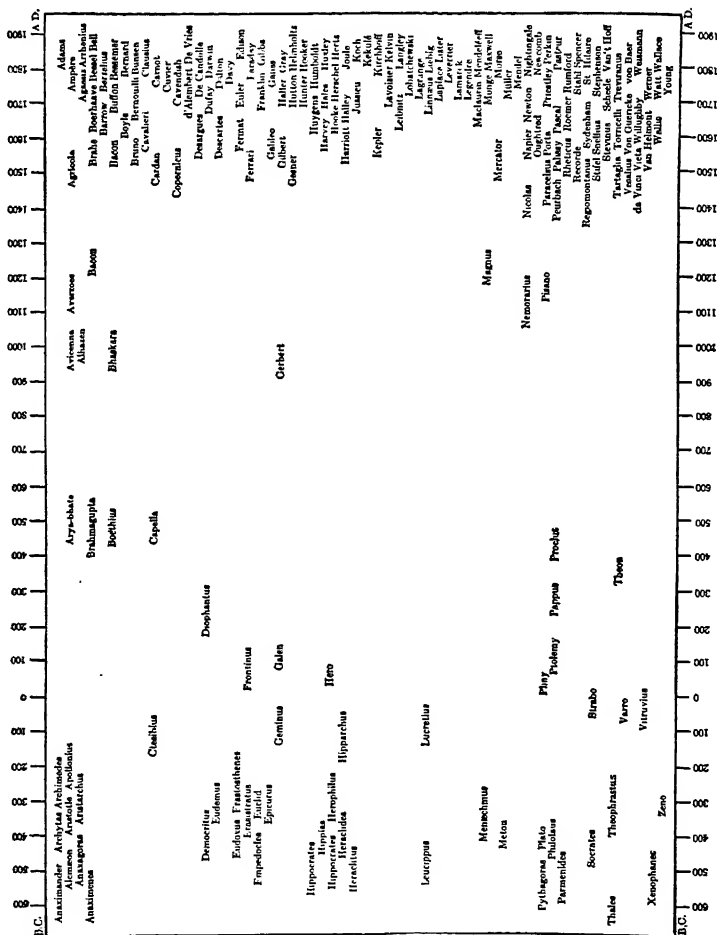
$|_p$  = 10 nations

*Comment on notation:*

The adpopulating processes in this aggregation of ten nations vary from about 50% to over 1500% in a century.

# A ROLL OF HONOR OF SCIENTISTS

This chart shows that the Greeks made their greatest contributions 600 B.C. to 300 A.C.; then come some Hindu, Alexandrian, and Roman names. The chart is blank for about 300 years. Then comes a rebirth of learning, but there are not many names for another 500 years. The last 200 years mark the great outburst of scientific knowledge.



*Ref.: Marshall, Leon O., The Story of Human Progress, Macmillan, 1925, p. 166.*

## S. 4 (Continued)

Descriptive formula:  $S_4 = {}^1T^{-1} : {}^pP$ 

Quantic number = 9;0;0;1

Legend:

 $S_4$  = The situation ${}^1|$  = beginning in 600 B.C.

records for each of

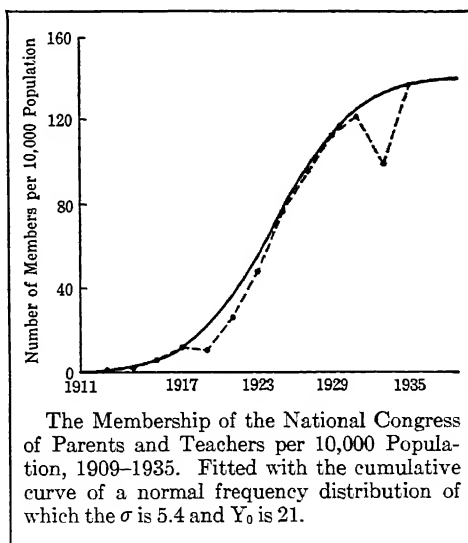
the corresponding

 ${}_t|$  = 25 periods ${}^pP$  = scientists listed by name $T^{-1}$  = of a century each

Comment on notation:

$P$  alone would denote a number of persons.  ${}^p|$  denotes an aggregation, listing individual persons,  $p$  in number.

## S. 5



Ref.: Pemberton, Earle H., "The Effect of a Social Crisis on the Curve of Diffusion," *American Sociological Review*, Vol. II, No. 1, Feb., 1937, p. 58.

Descriptive formula:  $S_5 = {}^1T^{-1} : {}^{\%}P_p$ 

Quantic number = 9;0;0;1

Legend:

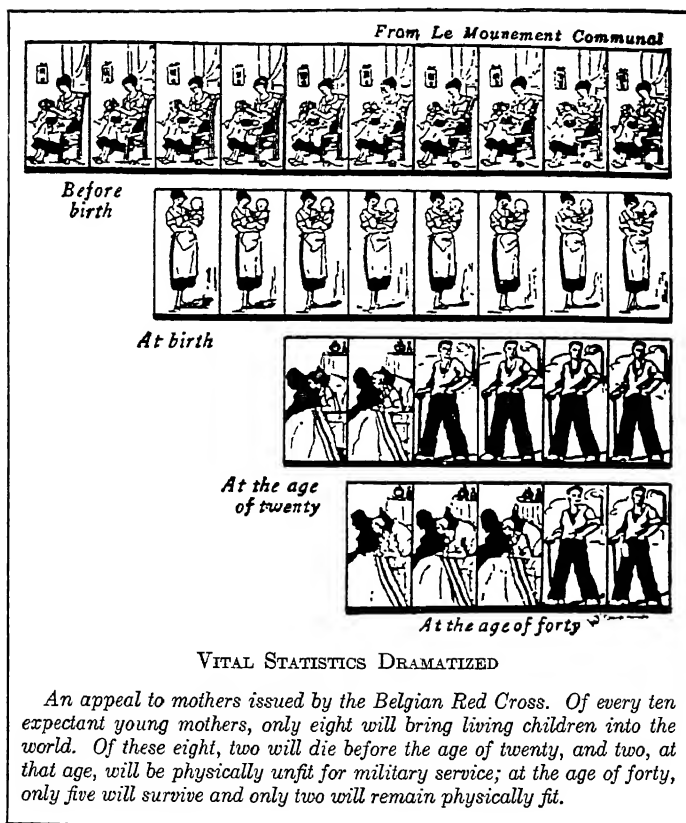
 $S_5$  = The situation ${}^{\%}P$  = the relative number of members of the N.C.P.T.

records for each of

for each of

 ${}_tT^{-1}$  = the 24 years $|_p$  = 2 plurels—actual and normal ${}^1|$  = beginning in 1911

## S. 6



Ref.: *The Survey*, Vol. XLVIII, No. 10, June 15, 1922, p. 401.

Descriptive formula:  $S_6 = \text{TT}^{-1} : \%P_p$

Quantic number = 9;0;0;1

Legend:

$S_6$  = The situation

: = subdivided into

records for each of

$|_p$  = the physically fit and unfit  
plurels

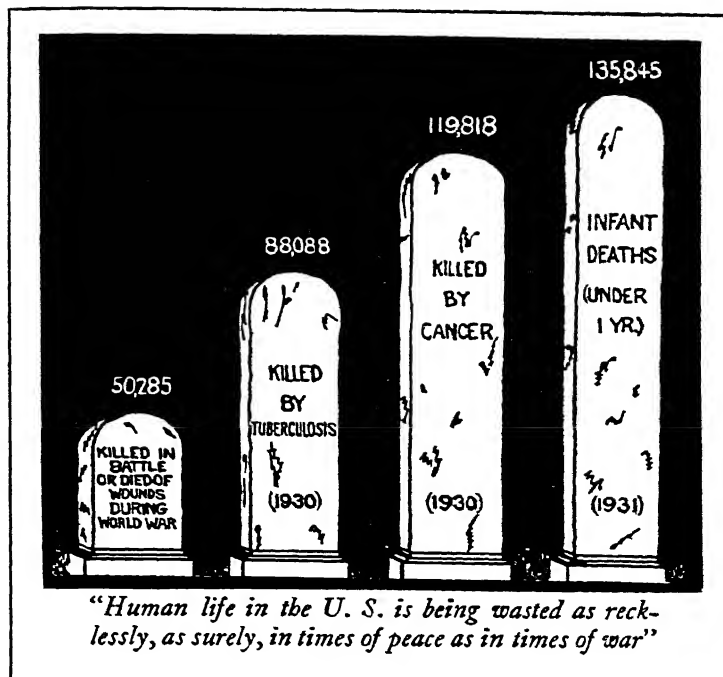
$\text{TT}^{-1}$  = 4 dates

$\%P$  = the percentage of Belgians

Comment:

The process of depopulating the plurel of the "fit" has gone on at an average velocity of 1.25% per year for the first 40 years.

## S. 7



Ref.: Emerson, Haven, "Medical Care for All of Us," *Survey*, Vol. LXVIII, No. 17, Dec. 1, 1932, p. 632.

Descriptive formula:  $S_7 = P_p : , T^{-1}$

Legend:

$S_7$  = The situation  
records

$P$  = the number of deaths  
in each of

Quantic number = 9;0;0;1

$|_p$  = 4 plurels (classified by cause)

$, T^{-1}$  = in corresponding specified periods

Comment on notation:

Note the unusual situation where the time period, being different for each plurel, is dependent on  $|_p$  and, therefore, follows the colon.

## S. 8

THE DISPLACEMENT OF LABOR BY INCREASING PRODUCTIVE  
EFFICIENCY AND ITS ABSORPTION BY AMERICAN INDUSTRY,  
1920-29

## MANUFACTURES

<i>Changes in Employment (+) or (-) during the Current Year</i>			
<i>Year</i>	<i>Due to Changes in Efficiency</i>	<i>Due to Changes in Output</i>	<i>Net Change since 1920</i>
1921	-163,000	-2,045,000	-2,208,000
1922	-935,000	+1,759,000	-1,384,000
1923	-183,000	+1,350,000	-217,000
1924	-276,000	-584,000	-1,077,000
1925	-495,000	+948,000	-624,000
1926	-93,000	+211,000	-506,000
1927	-68,000	-204,000	-778,000
1928	-503,000	+440,000	-841,000
1929	-116,000	+541,000	-416,000

## RAILROADS\*

<i>Changes in Employment (+) or (-) during the Current Year</i>		
<i>Due to Changes in Efficiency</i>	<i>Due to Changes in Output</i>	<i>Net Change since 1920</i>
+2,000	-494,000	-492,000
-36,000	+100,000	-428,000
-52,000	+286,000	-194,000
-47,000	-103,000	-344,000
-82,000	+80,000	-346,000
-39,000	+93,000	-292,000
+9,000	-67,000	-350,000
-74,000	-5,000	-429,000
-26,000	+39,000	-416,000

## COAL MINING †

<i>Changes in Employment (+) or (-) during the Current Year</i>			
<i>Year</i>	<i>Due to Changes in Efficiency</i>	<i>Due to Changes in Output</i>	<i>Net Change since 1920</i>
1921	-15,000	-165,000	-180,000
1922	-27,000	-62,000	-269,000
1923	-15,000	+224,000	-60,000
1924	+8,000	-94,000	-146,000
1925	-7,000	-19,000	-172,000
1926	+5,000	+102,000	-65,000
1927	-11,000	-66,000	-142,000
1928	-21,000	-25,000	-188,000
1929	-12,000	+29,000	-171,000

## TOTALS FOR THE 3 GROUPS

<i>Changes in Employment (+) or (-) during the Current Year</i>		
<i>Due to Changes in Efficiency</i>	<i>Due to Changes in Output</i>	<i>Net Change since 1920</i>
-176,000	-2,704,000	-2,880,000
-998,000	+1,797,000	-2,081,000
-250,000	+1,860,000	-471,000
-315,000	-782,000	-1,567,000
-584,000	+1,009,000	-1,142,000
-127,000	+406,000	-863,000
-70,000	-337,000	-1,270,000
-598,000	+410,000	-1,458,000
-154,000	+604,000	-1,003,000

\* Class I railroads.

† Anthracite and bituminous coal mining combined.

Source: David Weintraub, "The Displacement of Workers through Increases in Efficiency and Their Absorption by Industry," *Journal of the American Statistical Association*, December, 1932, pp. 396-97. The table covers wage-workers only.

Ref.: Corey, Lewis, *The Decline of American Capitalism*, Covici Friede, 1934, p. 227.

## S. 8 (Continued)

Descriptive formula:  $S_8 = tT^{-1} : P_p : a$ 

Quantic number = 9;0;0;1

Legend:

 $S_8$  = The situation

for each of

records for each of

 $|_p$  = 3 major industries and their sum, $tT^{-1}$  = the 9 years

subdivided into

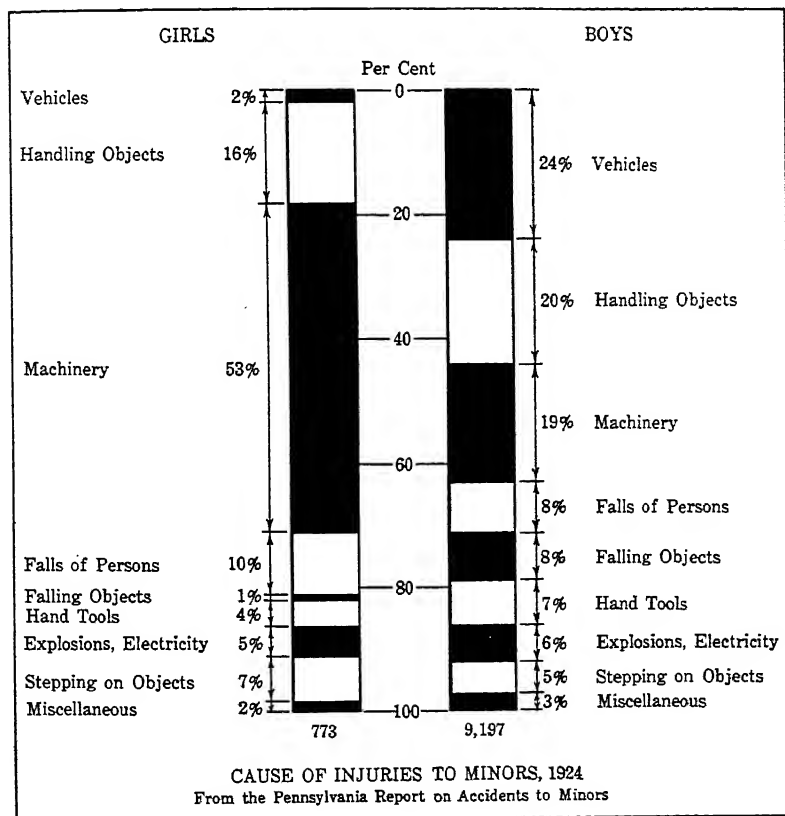
 $|$  = beginning in 1921

the changes in

 $|_a$  = 2 plurels identified by causes

P = the number of employees

## S. 9



S. 9 (*Continued*)

*Descriptive formula:*  $S_9 = {}^t({}_{\%}PT^{-1})_p : q$

*Quantic number* =  $9;0;0;1$

*Legend:*

$S_9$  = The situation

in each of

records

$|_p = 2$  sexes

${}_{\%}P$  = the percent of persons injured

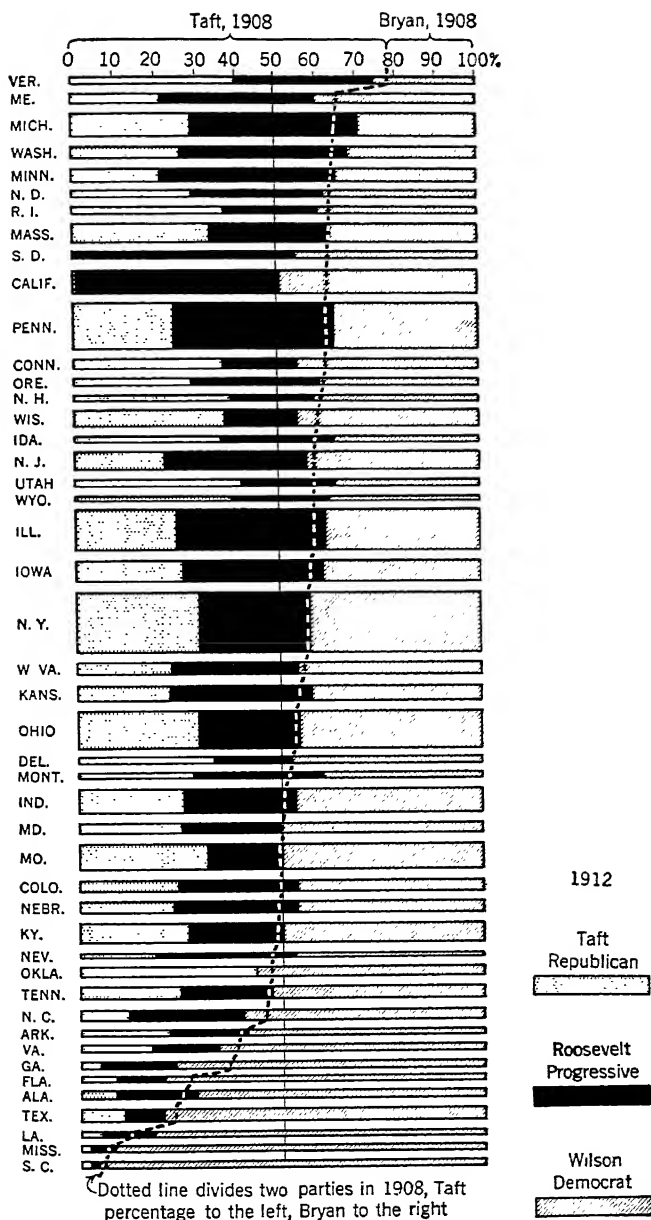
$:$  = subclassified by

$T^{-1}$  = annually

$|_q = 9$  causes

${}^t|$  = in 1924

# S. 10



THE VOTE FOR PRESIDENT IN 1908 AND IN 1912 BY STATES  
(Prof. Irving Fisher in the *New York Times*.)

Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Manufacturing Co., 1923, p. 10.

## S. 10 (Continued)

Descriptive formula:  $S_{10} = {}^tT^{-1} : {}^cP_p : q$ 

Quantic number = 9;0;0;1

Legend:

 $S_{10}$  = The situation ${}^cP$  = a % of U.S. voters

records

voting for each of

 ${}^tT^{-1}$  = on 2 dates (1908 and 1912) $|_q$  = the 3 major parties (2 in 1908)

for each of

 $|_p$  = 46 States

Comment:

In operating on the data presented in this third-degree matrix ( ${}^t| \times |_p \times |_q$ ) here, let the exponent, which specifies the operational thoroughness of studying the data, take successively the values of 0, 1, and 2.

First operate with a zero exponent:  ${}^tT^{-1} : ({}^cP)_q^0 = +{}_t|_q$  (in Brief-S), meaning that the aggregation of parties has increased from 2 to 3 in the four years, i.e., that a *dissociating* process has taken place to the extent of 50%. ( $|_p$  is a constant during this period and hence may be neglected in studying this change.)

Next with an exponent of 1, the processes of *depopulating* and *adpopulating* the Republican ( $|_{q'}$ ) and Progressive ( $|_{q''}$ ) parties respectively are illustrated:  $-{}_tP_{q'}$ ,  $+{}_tP_{q''}$ .

Next for secondary processes (exponent = 2), *mobility* of voters between political parties is illustrated:  $Mb_N = (.5 \sum_i {}^tP_i^2)^{-.5}$ . As an example, the mobility for the State of New York was:

	${}^cP$	${}^t{}^cP^2$	
Republicans	-20	400	Mb = 24%
Progressives	+27	729	
Democrats	-7	49	
	0	1178	

This 24% mobility means 24% of the maximal shift of votes, as when a monopoly by one party in 1908 becomes another's monopoly in 1912.

Thus a nullary process here measures the changing number of plurels; a primary process measures the changing size of each plurel taken by itself; and a secondary process measures the shifts between plurels in the total voting population.

As usual, election data are an example of our tension theory that people's desires for desiderata create tensions (see Ch. V). Votes may be considered either, (a) from the viewpoint of the voting population, as indices of the desire of the electorate, or (b) from the viewpoint of the candidates, as the desideratum desired by the candidates:  ${}^t(PD = VE)_p^s : q$ .

Since in a democracy each voter's intensity of desire is considered equal to any other voter's and hence is the unit of desire,  $D = 1$ . The desideratum to the voters is a particular candidate and hence is an all-or-none quantity, so that  $V^e = V^0 = 1$ . Hence  $P = E$ , the number of voters (P) voting for a candidate measures the tension (E) of the population towards him.

Again operate on the data as specified by exponents of 0, 1, and 2 to measure processes defined by the tension theory. First, from the point of view of the voters as the population, P, the number of candidates is the number of unitary qualitative values,  $V_1^0 = 2$  in 1908, and 3 in 1912. Thus the process of *dissimilarizing*, which was defined as increasing the number of qualitative characteristics in a situation, has gone on to the extent of 50% ( $+ \frac{1}{2} = 50\%$ ). This process is identical with dissociating in this situation, since a political party (a plural) and a candidate (a desideratum) are interchangeable in the data as presented.

Next, with exponents of 1 on D, those voters who change their party in 1912 illustrate *devaluating* the desideratum represented by their former party, and *evaluating* the desideratum represented by their new party. Thus, Republicans who became Progressives devaluated, i.e., decreased their intensity of desire for, the Republican candidate and evaluated, i.e., increased their intensity of desire for, the Progressive candidate.

Next, with exponents of 2 on D the process of *revaluating*, Rl, i.e., the shifting of valuations, D, is calculable by Eq. 54, Ch. X. Since D is 1 for the voters who vote for a certain candidate and zero for those voters towards every other candidate, the shift of desire between candidates is identical with the shift of voters between parties, and so the revaluating, Rl, becomes identical with the mobility, Mb, i.e., 24% in the case of New York.

Second, from the point of view of the candidates (or the political parties) the population is defined as  $l_0$  in this situation, and is 2 in 1908, and 3 in 1912; D is now defined as the intensity of their desire to win the election and is unmeasured in the data as presented; and V is now defined as the votes, since these are the desideratum desired by the candidates, so that now  $V = P$  (if D is roughly assumed to have been constant). When the exponent is 1, the process of *progressing*, i.e., increase of the desideratum (votes), has taken place for the Progressive Party, and *regressing*, i.e., decrease of votes, has taken place for the Republican Party.

Next, when the exponent is 2, the process of *competing* between parties is measured. The amount of competing, Cp, is the same as the mobility, Mb, since P, the number of voters, is here V, the amount of the desideratum. Thus for New York there was 24% of maximal competing; for California 58%; for South Carolina 5%. (These figures are approximate ones from a visual reading of the graph.)

In the above defined situations, the resultant processes of *tensing* and *retensing* are measurable wherever P, D, and V are simultaneously determinable.

These illustrations of the processes defined by the tension theory emphasize the necessity in applying this theory to specify rigorously the desideratum, the population desiring it, and the index of their intensity of desire, or else the tension theory equation is indeterminate, and reasoning in terms of its factors may be confusing.

## S. 11

## CAUSAL FACTOR ANALYSIS

Comparison of Statistics for the Year 1923-24 of the Chief Family Helping Agencies (three private, one public) with the Boston Provident Association's Figures for 1913

	<i>Overseers of Public Welfare</i>	<i>Provident Associa- tion *</i>	<i>Family Welfare Society</i>	<i>Federated Jewish Charities</i>	<i>Total</i>	<i>1913 Prov. Assoc.</i>
<i>Total No. of Case Units</i>	4680	837	3315	2011	10843	875
<i>Industrial Accident Cases</i>	50	53	96	19	218	103
<i>Percent</i>	1.06%	6.3%	2.8%	1.0%	2.01%	13%
<i>Intemperance Cases</i>	88	84	270	0	442	180
<i>Percent</i>	1.8%	9.9%†	8.0%	0	4.0%	20%
<i>Unemployment Cases</i>	690	370	1066	435	2561	245
<i>Percent</i>	14.7%	44.2%	32.0%	21.0%	23.6%	28%
<i>Illness Cases</i>	665	331	2108	929	4033	269
<i>Percent</i>	14.2%	39.5%	63.5%	46.2%	37.1%	30%
<i>Tuberculosis Cases</i>	148	39	128	112	427	129
<i>Percent</i>	3.1%	4.6%	3.8%	5.5%	3.8%	14%
<i>Desertion and Non-support Cases</i>	512	97	331	168	1408	135
<i>Percent</i>	10.9%	11.5%	10.0%	8.3%	10.2%	15%

\* Intake only.

† Rate declining the past 7 months to 7.5%.

Note the size and persistency of the "Illness" factor and of "Unemployment."

Note the substantial reduction of the "Tuberculosis" factor which may be set alongside the recorded decline in our tuberculosis morbidity figure which, in the same period, dropped 4.45%.

Note that "Intemperance," which dropped from 20% to 1% in the Provident Association analysis in 1919, then rose to 9.9% last year and is now falling to 7.5%.

Note that "Industrial Accident" appears to but one-half the extent it did in 1913.

Ref.: Davis, Jerome and Barnes, Harry E., *Readings in Sociology*, D. C. Heath and Co., 1927. p. 907.

*Descriptive formula:*  $S_{11} = \text{†T}^{-1} : \% . P_p : q$

*Quantic number* = 9;0;0;1

*Legend:*

$S_{11}$  = The situation  
records for each of

$\% . P$  = numbers and  $\%$ 's of cases  
of

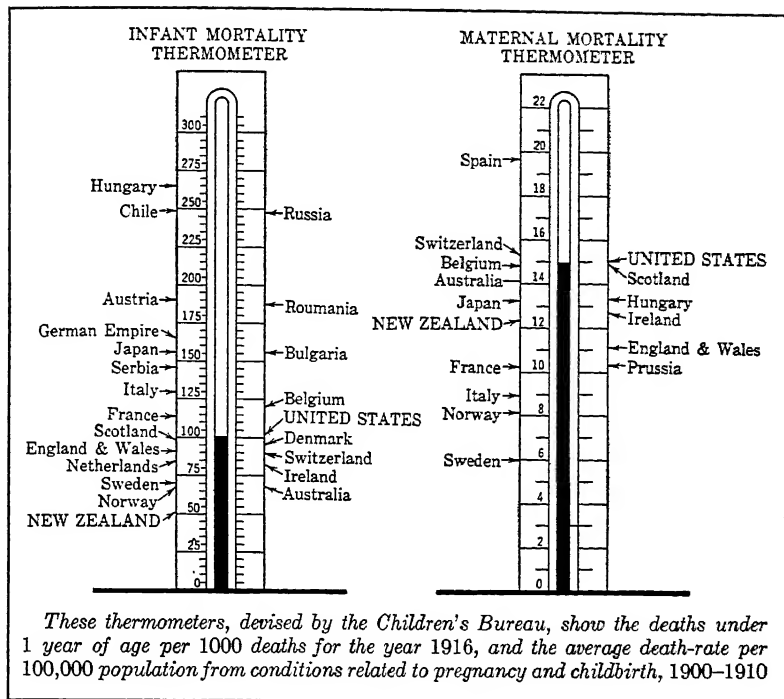
$\text{†T}^{-1}$  = 2 particular periods: 1913,  
1923-24

$|_p$  = 6 causal types and their sum  
subclassified into

$:$  = the corresponding

$|_q$  = 4 agencies

## S. 12



Ref.: Editorial, "Red Tape Made Colorful," *Surrey*, Vol. LXIII, No. 10, February 15, 1930, p. 603.

Descriptive formula:  $S_{12} = (\%P_p : vT^{-1})_q$   
Legend:

$S_{12}$  = The situation  
records

$\%P$  = the death rate  
 $|_p$  = in mother and infant plurals

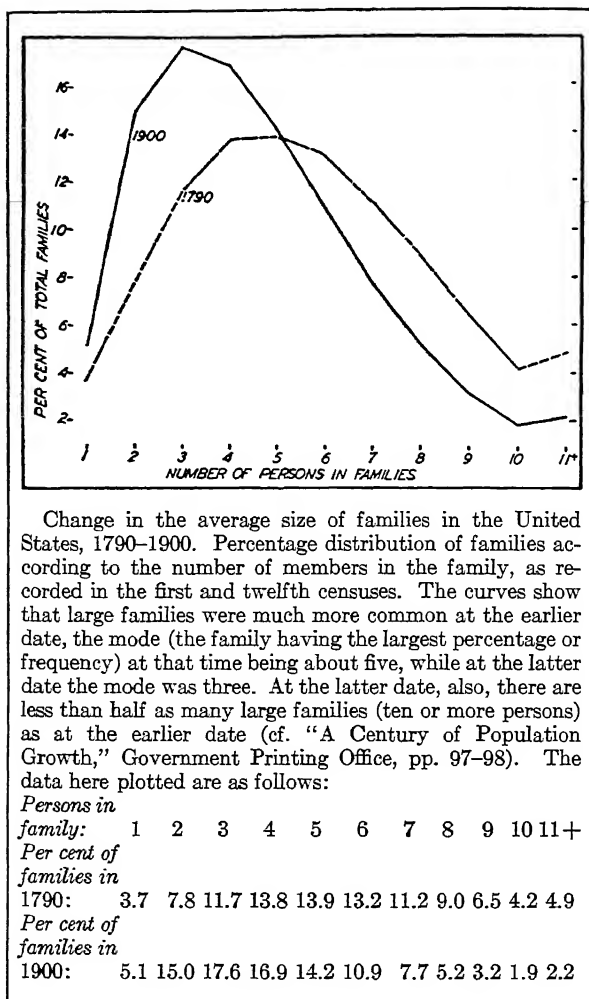
Quantic number = 9;0;0;1

$vT^{-1}$  = for corresponding specified  
periods

in each of

$|_q$  = 23 countries

## S. 13



Ref.: Reinhardt and Davies, *Principles and Methods of Sociology*, Prentice-Hall, 1932, p. 484.

S. 13 (*Continued*)

*Descriptive formula:*  $S_{13} = {}^tT^{-1} : {}_p : \Sigma_q P$

*Quantic number* = 9 ; 0 ; 0 ; 1

*Legend:*

$S_{13}$  = The situation

: = with the corresponding

records for each of

$\Sigma_q$  | = number of families of each  
size

${}^tT^{-1}$  = 2 dates, 1790, 1900

${}_pP$  = the number of persons per  
family, i.e., 11 size-plurels

*Comment on notation:*

The populational class-interval scripts illustrate a homosectoral frequency distribution, here a distribution of plurels-by-size.



S. 14 (*Continued*)

*Descriptive formula:*  $S_{14} = 'T^{-1} : ^P P : I^0$

*Quantic number* = 9;0;0;1

*Legend:*

$S_{14}$  = The situation

together with

records for each of

$^P|$  = 3 typical persons

${}_t T^{-1}$  = 18 years (abscissae)

and

$'|$  = beginning in 1919

$I^0$  = their emotional reactions

P = the number of Germans unemployed (ordinates)

ORIENTAL LEADERSHIP		EUROPEAN LEADERSHIP	
8000 B. C.		1200 B. C.	
STONE AGE IN THE ORIENT		COPPER-BRONZE AGE	
NEOLITHIC		IRON AGE	
<p><b>PALEOLITHIC</b></p> <p>Western Asia retarded by Cold</p> <p>Domesticated Animals</p> <p>Domesticated Grains</p> <p>Pottery</p> <p>Transition from Hunting to Agricultural Stage</p>	<p>Earliest Civilization</p> <p>Egyptian Ships in Mediterranean</p> <p>Egyptian Civilization reaches Southeastern Europe by Ship.</p>	<p>Babylonian Commerce through Asia</p> <p>Babylonian and Egyptian Civilization reach Southern Europe through Asia Minor and Phoenicians.</p> <p>Phoenician Commerce</p>	<p>Modern Nations</p> <p>Middle Ages</p>
<p>Europe retarded by ice</p> <p>Europe free from ice and cold</p> <p>First Advance of Europe under Oriental Influences.</p> <p>Pottery (?)</p> <p>Agriculture</p> <p>Domestic Animals</p>		<p>Third Advance of Europe under Oriental Influences.</p> <p>Europe gains Leadership of the World</p> <p>Coinage, Bronze, Casting, Colonnade, Alphabet, Glass, Textiles, Ships, etc.</p> <p>Class, Writing, Barbarian Invasions</p> <p>Navigation, Metal, Pottery, Wheel, and Furnace, etc.</p> <p>Second Advance of Europe under Oriental Influences.</p> <p>Earliest Civilization in Europe</p>	
<p><b>PALEOLITHIC</b></p> <p>STONE AGE IN EUROPE</p>		<p><b>BRONZE AGE</b></p>	
<p><b>NEOLITHIC</b></p>		<p><b>IRON AGE</b></p>	

KEY ——— Orient ——— Europe

*Ref.:* Davis, J. and Barnes, H. E., *An Introduction to Sociology*, D. C. Heath, 1927, p. 66.

S. 15 (*Continued*)

*Descriptive formula:*  $S_{15} = {}_t : {}_u T^{-1} : \underline{P}_p : I_1^0$

*Quantic number* = 9;0;0;1

*Legend:*

$S_{15}$  = The situation  
records

$\underline{P}_p$  = the dominant peoples  
with their corresponding

${}_t T^{-1}$  = for a series of ages

$I_1^0$  = achievements

${}_u |$  = and subperiods (paleolithic,  
etc.)

$|_t$  = in 2 chronologies (Europe and  
the Orient)

*Comment:*

This situation records processes of differing,  ${}_t I^0$  (Eq. 25a, Ch. X), as qualitative implements, activities, and peoples are replaced by later ones; and roughly suggests the process of dissimilarizing  ${}_t |_{\Sigma_1}$  (Eq. 26, Ch. X), as the number of such qualitative entities increases with invention.

S. 16

**The Family's  
Vacation "Musts"**

Dad		Fishing—Golf—Camping Mountain Air and Sunshine
Mother		Fine Living—Horseback Riding—Parties—Rest
Sis		Swimming—Tennis Dancing—Hiking
Bill		Camera Hunting—Pack Trips—Mountain Climbing
Aunt Lil		Driving—Sight Seeing Bridge—Good Food
Bud		Slides—Swings—Sandbox Nursery games—Pony Arena

**It all adds up To  
Sun Valley!**

ANOTHER TRIUMPH OF THE PROGRESSIVE  
UNION PACIFIC

Ref.: *Time*, April 25, 1938, p. 33.Descriptive formula:  $S_{16} = {}^n P : (I^0 T^0 \cdot -1)_i$ 

Quantic number = 9;0;0;1

Legend:

 $S_{16}$  = The situation

of

records for each of

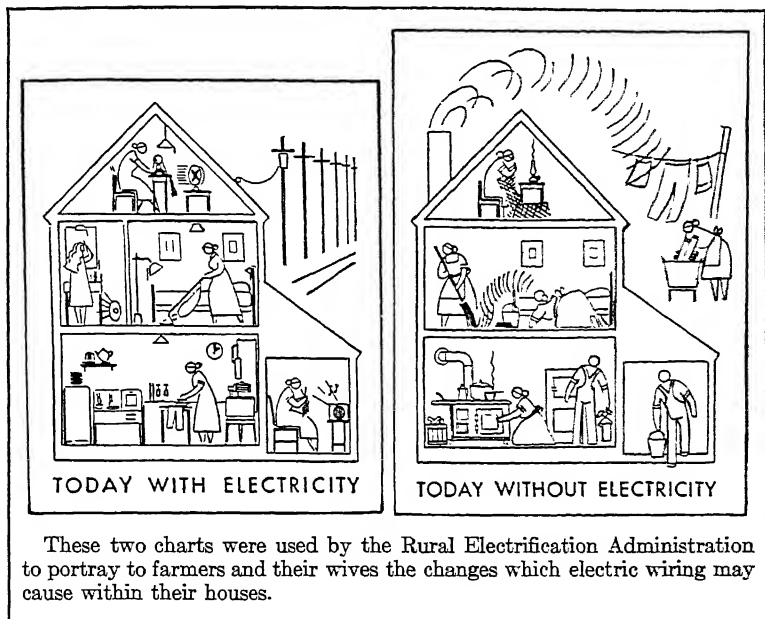
 $I^0 T^0$  = qualitative states ${}^n P$  = 6 persons

or

 $|_i$  = a series $I^0 T^{-1}$  = activities*Comment on notation:*

This situation illustrates the use of the attribute symbol,  $I^0$ , divided by time to denote a velocity of a qualitative change, i.e., an event or activity unspecified in amount.

## S. 17



These two charts were used by the Rural Electrification Administration to portray to farmers and their wives the changes which electric wiring may cause within their houses.

Ref.: Modley, Rudolf, *How to Use Pictorial Statistics*, Harpers, 1937, pp. 114-115.

Descriptive formula:  $S_{17} = {}^pP : (I^0T^{-1})_{i,j}$

Legend:

$S_{17}$  = The situation  
records with respect to

${}^pP$  = 2 houseworkers

$(I^0T^{-1})$  = activities

in each of

Quantic formula = 9;0;0;1

$|_i$  = 2 classes (with and without electricity)

of

$|_j$  = 12 kinds

## S. 18

<i>Question</i>	<i>Before</i>	<i>After</i>	<i>Gain</i>	<i>Loss</i>
1. Do you believe in God as a supreme, all powerful, all knowing, personal being who holds direct communication with man?	<i>yes</i> 49 <i>no</i> 17	<i>yes</i> 51 <i>no</i> 24	2  7	
2. Do you believe that God created the world and all contained therein in the manner described in the first chapter of Genesis?	<i>yes</i> 45 <i>no</i> 24	<i>yes</i> 27 <i>no</i> 46	  22	18
3. Do you believe that man was created by a single act and in the image of God as described in Genesis?	<i>yes</i> 44 <i>no</i> 24	<i>yes</i> 16 <i>no</i> 58	  34	28
4. Do you believe that the Bible is verbally inspired?	<i>yes</i> 25 <i>no</i> 38	<i>yes</i> 14 <i>no</i> 53	  15	11
5. Do you believe that Christ rose from the dead on the third day after his burial leaving the tomb empty?	<i>yes</i> 49 <i>no</i> 18	<i>yes</i> 53 <i>no</i> 17	4   	1
6. Do you believe that Jesus Christ was "the only begotten Son of God" born of a virgin and without a human father?	<i>yes</i> 43 <i>no</i> 23	<i>yes</i> 31 <i>no</i> 37	  14	12
7. Do you believe in miracles as events brought about through divine interference with the laws of nature?	<i>yes</i> 46 <i>no</i> 21	<i>yes</i> 19 <i>no</i> 51	  30	27
8. Do you believe in a Heaven and a Hell as definite places where the righteous and wicked will spend eternity?	<i>yes</i> 26 <i>no</i> 39	<i>yes</i> 5 <i>no</i> 68	  29	21
9. Do you believe in prayer as a means of altering the regular operation of natural law?	<i>yes</i> 17 <i>no</i> 47	<i>yes</i> 10 <i>no</i> 58	  11	7
TOTAL	<i>yes</i> 344 <i>no</i> 251	<i>yes</i> 226 <i>no</i> 412	  161	118
TOTAL GAIN	<i>Yes</i>		6	124
TOTAL LOSS	<i>No</i>		162	1

Ref.: Binnewies, W. G., "Measuring Changes in Opinion," *Sociology and Social Research*, Vol. XVI, Sept., 1931, pp. 144 and 145.

## S. 18 (Continued)

Descriptive formula:  $S_{18} = {}^tT^{-1} : I_1^0 : P$ 

Quantic number = 9;0;0;1

Legend:

 $S_{18}$  = The situation $I_1^0$  = 9 attributes (of belief)

records for each of

 $P$  = a frequency of persons endorsing ${}^t$  = 2 dates ${}^tT : P$  = including gain or loss of endorsers

and for

 ${}^tT^{-1}$  = their unspecified interval

for each of

## Comment:

As usual in situations involving votes, attitude tests, etc., these can be considered as indices of desire,  $D$ , in the tension theory. In this situation, each vote (i.e., endorser) is one unit of the observed population's intensity of desire,  $D$ , for that qualitative desideratum (belief),  $I^0 \equiv V^0$ . Beliefs showing increase of endorsers are *evaluated*,  $+D$ , while beliefs with decreasing endorsers are *devaluated*,  $-D$ . The shifts between the nine beliefs is the process of *revaluating* a collection of desiderata and is measured by  $Rl$  (Eq. 53, Ch. X). For the nine beliefs together by the indicator of affirmation ("yes")  $Rl = 35\%$  and by the indicator of denial ("no")  $39\%$ . The situation excerpted omits those who were neutral or had no opinion. Averaging the two indicators, yields  $37\%$  of maximal revaluating recorded in this situation of religious beliefs.

$${}^t\%(PD \text{ or } \Sigma D) \quad {}^t\%(PD)^2$$

$$\left( \frac{118}{334} = 35\% \quad \begin{array}{l} \text{Affirmers} \\ \text{Non-affirmers} \end{array} \quad \begin{array}{r} - 35\% \\ + 35\% \\ 0 \end{array} \quad \frac{1225}{2450} \right) \times .5 = 1225 \sqrt{1225} = 35\% = Rl$$

Since the desideratum is a unitary one,  $V^0 = 1$ , the tension is also measured by the number of votes  ${}^t(PD = V^0E = P = E)$ , so that the process of *retensing*,  $Rt$ , (Eq. 57d, Ch. X) has the same index as revaluating. Also by considering voters as changing from one belief-plurel to another, *mobility* is measured by the same index,  $Mb = Rt = Rl$  here.

## S. 19

---

 ONE CENTURY COMPARED WITH ALL  
PRECEDING CENTURIES
 

---

*Some Steps in Progress in the  
Nineteenth Century*

1. Railways
2. Steamships
3. Electric telegraphs
4. Telephone
5. Matches
6. Gas illumination
7. Electric lighting
8. Photography
9. The phonograph
10. X-rays
11. Spectrum analysis
12. Anaesthetics
13. Antiseptic surgery
14. Principle of conservation of energy established
15. Molecular theory of gases
16. Velocity of light directly measured and earth's rotation experimentally shown
17. The discovery of the uses of dust
18. Chemistry, definite proportions
19. Meteors and the meteoritic theory
20. The proof of glacial epochs
21. The proof of the antiquity of man
22. Organic evolution established
23. Cell theory and embryology
24. Germ theory of disease

*Some Steps in Progress in All  
Preceding Ages*

1. The mariner's compass
  2. The steam engine
  3. The telescope
  4. The barometer and thermometer
  5. Printing
  6. Arabic numerals
  7. Alphabetical writing
  8. Modern chemistry founded
  9. Electric science founded
  10. Gravitation established
  11. Kepler's laws on the Motion of Planets
  12. The differential calculus
  13. The circulation of the blood discovered
  14. Light proved to have finite velocity
  15. The development of geometry
  16. Gunpowder
  17. Paper
  18. Fire making
  19. Tool making
  20. Agriculture
  21. Domestication of animals
  22. Metals and pottery
- 

Ref.: Marshall, Leon C., *The Story of Human Progress*, Macmillan, 1925, p. 169.

*Descriptive formula:*  $S_{19} = \text{tT}^{-1} : I_1^0$

*Legend:*

$S_{19}$  = The situation  
records for each of

$\text{tT}^{-1}$  = 2 periods (19th century and  
all previous centuries)

*Quantic number* = 9;0;0;0

$I_1^0$  = qualitative characteristics of  
progress, 46 in all

*Comment:*

The process of dissimilarizing  $+\text{t}|_{\Sigma 1}$ , is illustrated, and also the process of progressing for those plurels which desire these "steps," so that the number of steps becomes the quantity of the desideratum  $+\text{t}|_{\Sigma 1} = +\text{t}V$ , in the tension theory.

## S. 20

		EGYPTIAN	GREEK	LATIN	
1	Eagle . .	𐦃	Α Α α	Α Α α α α	Α
2	Crane . .	𐦃	Β Β β	Β Β β β	Β
3	Throne . .	𐦃	Γ Γ γ	Γ Γ γ γ γ	Γ
4	Hand . .	𐦃	Δ Δ δ	Δ Δ δ δ δ	Δ
5	Mæander . .	𐦃	Ε Ε ε	Ε Ε ε ε ε	Ε
6	Cerastes . .	𐦃	Ζ Ζ ζ	Ζ Ζ ζ ζ ζ	Ζ
7	Duck . .	𐦃	Η Η η	Η Η η η η	Η
8	Sieve . .	𐦃	Θ Θ θ	Θ Θ θ θ θ	Θ
9	Tongs . .	𐦃	Ι Ι ι	Ι Ι ι ι ι	Ι
10	Parallels . .	𐦃	Κ Κ κ	Κ Κ κ κ κ	Κ
11	Bowl . .	𐦃	Λ Λ λ	Λ Λ λ λ λ	Λ
12	Lioness . .	𐦃	Μ Μ μ	Μ Μ μ μ μ	Μ
13	Owl . .	𐦃	Ν Ν ν	Ν Ν ν ν ν	Ν
14	Water . .	𐦃	Ξ Ξ ξ	Ξ Ξ ξ ξ ξ	Ξ
15	Chair-back . .	𐦃	Ο Ο ο	Ο Ο ο ο ο	Ο
16	....	𐦃	Π Π π	Π Π π π π	Π
17	Shutter . .	𐦃	Ρ Ρ ρ	Ρ Ρ ρ ρ ρ	Ρ
18	Snake . .	𐦃	Σ Σ σ	Σ Σ σ σ σ	Σ
19	Angle . .	𐦃	Τ Τ τ	Τ Τ τ τ τ	Τ
20	Mouth . .	𐦃	Υ Υ υ	Υ Υ υ υ υ	Υ
21	Inundated Gard-n	𐦃	Φ Φ φ	Φ Φ φ φ φ	Φ
22	Lasso . .	𐦃	Χ Χ χ	Χ Χ χ χ χ	Χ

## HISTORY OF SOME OF OUR LETTERS

What ones can easily be traced back to Egypt?

Ref.: Marshall, Leon C., *The Story of Human Progress*, Macmillan, 1925, p. 223. From Clodd: *The Story of the Alphabet*, D. Appleton and Co.

Descriptive formula:  $S_{20} = T^{-1} : u : I^0$

Quantic number = 9;0;0;0

Legend:

$S_{20}$  = The situation  
records for each of  
 $|_t$  = 5 qualitatively characterized  
periods

$T^{-1}$  = of indefinite duration  
and their

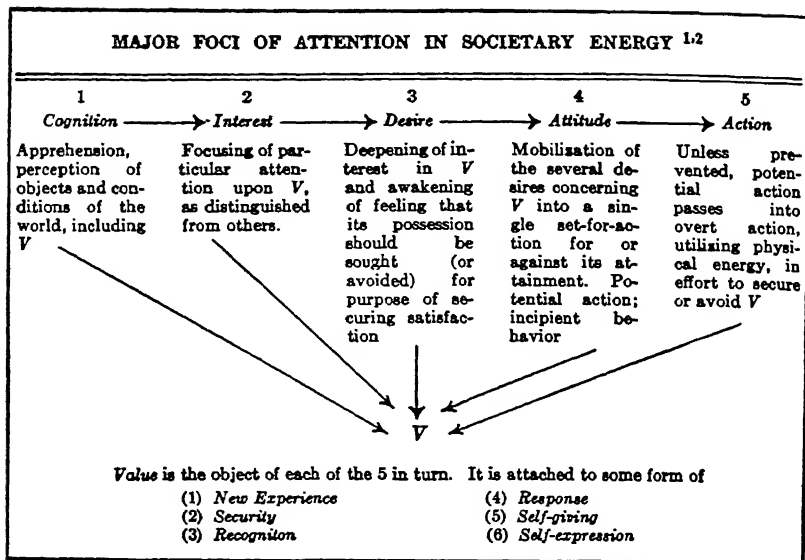
$|_u$  = several subperiods  
: = the corresponding  
 $I^0$  = shape, or symbol,  
of

$|_i$  = the 22 letters of the alphabet

Comment:

The process here illustrated is an aggregated differing ( $I^0$ )<sub>i</sub> (Eq. 25, Ch. X), the changing of the qualitative nature of each of a collection of characteristics. (Although an aggregated process, since the aggregated indices are homogeneous, it is nearer to the single process of "differing" than to the heterogeneous aggregated processes discussed in section IV of Chapter X.)

## S. 21



Ref.: Eubank, Earle E., *The Concepts of Sociology*, D. C. Heath, 1931, p. 202.

*Descriptive formula:*  $S_{21} = {}^tT^{-1} : I^0 : ;$

*Quantic number* = 9;0;0;0

*Legend:*

$S_{21}$  = The situation  
records for

$I^0$  = a "value," V  
in

$T^{-1}$  = an indefinite period  
in

$|$  = 6 subforms ("new experience" to "self-expression")

${}^t|$  = 5 successive stages

*Comment:*

The successive psychological states are an example of a series of qualitative changes which is defined as the process of differing ( ${}_iI^0$ ) in Eq. 25a, Ch. X.

## S. 22

TECHNIQUES OF CONTROL, ARRANGED ACCORDING TO THE  
ASCENDING LEVELS OF INTERSTIMULATION

*Form of Stimulus* (from the controller)      *Form of Response* (by the one controlled)

## I. General

1. General previous con-  
conditioning

## II. Specific

2. Stimulus of mere pres-  
ence  
3. Suggestion  
(a) By setting a copy  
(b) By initiating a  
thought

These may be conscious  
or unconscious, and either  
intentional or uninten-  
tional, on either side

4. Persuasion  
(a) By appeal to emo-  
tion  
(b) By appeal to reason  
5. Coercion  
(a) Non-violent, using  
psychological pres-  
sure  
(b) Violent, using physi-  
cal force

These can only be both  
conscious and intentional  
on both sides

## I. General

1. Response to general  
previous conditioning

## II. Specific

2. Spontaneous Reaction  
to mere presence  
3. Imitation  
(a) Mimetic (replica-  
tion)  
(b) Non-mimetic

4. Concurrence  
(a) Emotional  
(b) Rational  
5. Acquiescence  
(a) Compliance  
(b) Submission

Ref.: Eubank Earle Edward, *The Concepts of Sociology*, D. C. Heath and Co., 1931, p. 234.

*Descriptive formula*:  $S_{22} = {}_1T^{-1} :: I_1^0 : j : k, 1$

*Quantic number* = 9;0;0;0

*Legend*:

$S_{22}$  = The situation  
records

and

${}_1T^{-1}$  = 2 stages (stimulus and re-  
sponse)

$|_j$  = 5 minor classes

and

$::$  = cross-classified with

$|_k$  = 6 subclasses

and

$I_1^0$  = qualitative characteristics of  
interstimulation

$|_{,1}$  = 2 all-or-none other attributes  
are also noted ("conscious"  
and "intentional")

subclassified into

$|_1$  = 2 major classes       $\left\{ \begin{array}{l} \text{general} \\ \text{and} \\ \text{specific} \end{array} \right.$

*Comment*:

This situation is on the borderline of being the "quantitatively recorded societal situation" which alone S-theory claims to analyze. Its tabular form is an irregular matrix with qualitative verbal cell entries. As these change in time from the stimulus to the response stages, the S-theory aggregated process of differing  $({}_1I^0)_i$  is illustrated. (Eq. 25a, Ch. X.)

## S. 23

	Number of Styles	
	Actual	Proposed Simplification
Vitrified paving brick . . . . .	66	5
Metal lath . . . . .	125	24
Woven wire fencing . . . . .	552	69
Woven wire fence packages . . . . .	2,072	138
Asphalt penetrations . . . . .	88	9
Roofing slate . . . . .	60	30
Hollow tile . . . . .	36	19
Rough face brick . . . . .	39	1
Smooth face brick . . . . .	36	1
Files and rasps . . . . .	1,351	496
Range boilers . . . . .	130	13
Beds, springs and mattresses . . . . .	78	4
Bed blankets—sizes . . . . .	78	12
Forged tools . . . . .	665	351
Blackboard slate . . . . .	90% elimination	endorsed
Bolts and nuts (plows) . . . . .	40% “	“

Ref.: Chase, Stuart, *The Tragedy of Waste*, Macmillan, 1925, pp. 168-169.

Descriptive formula:  $S_{23} = {}^t\underline{T}^{-1} : I_1^0 : \Sigma_j$

Quantic number = 9;1;0;0

Legend:

$S_{23}$ = The situation records	$ _i$ = 16 types : = each subclassified
${}^t $ = on 2 dates	into
$\underline{T}^{-1}$ = with an unspecified time interval	$ \Sigma_j$ = a number of styles
$I^0$ = manufactured articles of	

Comment:

This shows that the process of assimilizing,  ${}^tI_{\Sigma_j}^0$  (Eq. 26, Ch. X), is recorded in each of an aggregation of 16 manufactured articles.

The operation of summing attributes, i.e. counting the number of them, converts the aggregated nullary indicator into a primary indicator with a quantic digit of unity.

## S. 24

# The Appropriation Pie



(Courtesy of Labor)

## United States Appropriations, 1920

I. Past Wars	\$3,855,482,586	68%
II. Future Wars	1,424,138,677	25%
III. Civil Departments	181,087,225	3%
IV. Public Works	168,203,557	3%
V. Research, Education and Health	57,093,661	1%
Total	\$5,686,005,706	100%

Ref.: Card No. 5 of the "Facts on Disarmament" Exhibit, Disarmament Education Committee, 629G St., N. W., Washington, D. C.

Descriptive formula:  $S_{24} = t'(\$. \% IT^{-1})_i$

Legend:

$S_{24}$  = The situation  
records

$t'$  = in 1920

$T^{-1}$  = the annual

$\$. \% I$  = appropriations (in \$ and %  
units)

Quantic number = 9;1;0;0

for

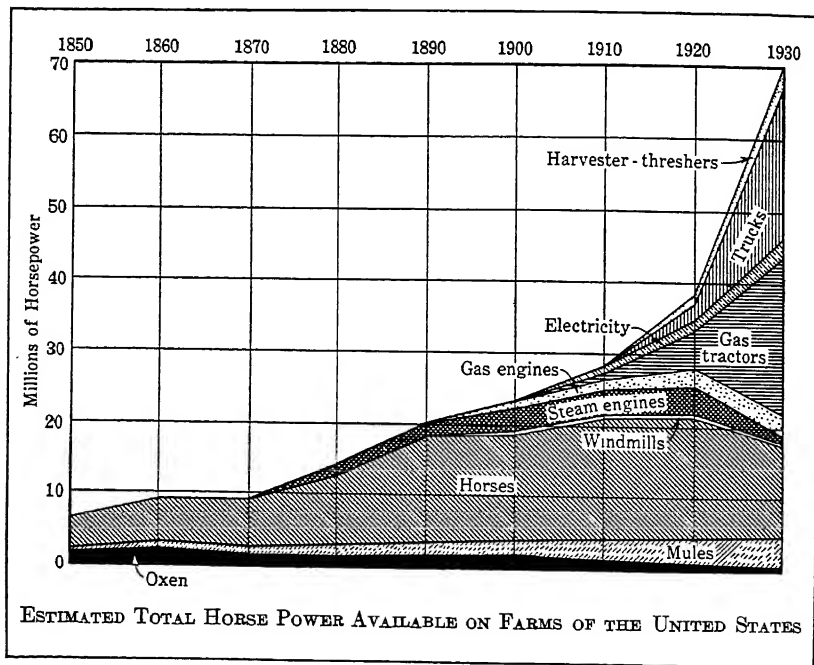
$|_i$  = 5 kinds of Government expenditure

## S. 24 (Continued)

## Comment:

The process of effective competing between the Federal Government's functions is illustrated. The index of competing,  $C_p$  (Eq. 47a, Ch. X) is 40%, where with five competitors starting equally from zero appropriation at the beginning of the year the "roof," or maximal,  $C_p$  is 63%. This means 40/63, or 63.5% of a complete monopoly of Federal expenditures by one of the five functions listed.

## S. 25



Ref.: President's Research Committee *Social Trends*, McGraw-Hill Book Co., 1933, p. 101.

Descriptive formula:  $S_{25} = {}^{\frac{1}{4}}T^{-1} : I_1$

Quantic number = 9;1;0;0

Legend:

$S_{25}$  = The situation  
records for

$\frac{1}{4}$  = beginning in 1850

$I$  = an indicator of horsepower  
for each of

$\frac{1}{4}$  = 8

$T^{-1}$  = decades

$|_1$  = 10 prime movers

## Comment:

This S-situation illustrates three processes as defined by the V-theory.

1. "Progressing,"  $+V$

Assuming that horsepower is a desideratum, i.e., an indicant desired by

farmers,  $I \equiv V$ , gives  $+_tV$  (the quantity of the desideratum is increasing with time). As this is due to human effort, it is  $V_{co}$ , "co-operating."

"Progressing" or "Co-operating"

$$= +_tV_{co} = \frac{100(1930_{Hp.} - 1850_{Hp.})}{1850_{Hp.}} = \% \text{ increase}$$

$$= \frac{100(70 \times 10^6 - 7 \times 10^6)}{7 \times 10^6} = 900\% \text{ in 80 years}$$

This is an average annual velocity of the process of co-operating to increase farm horsepower of 11.25% of the amount in 1850.

## 2. Dissimilarizing, $+_t\Sigma$

As the number of sources of horsepower increases from 3 kinds in 1850 to 9 kinds in 1930, "dissimilarizing" has gone on, and its amount is  $+_t\Sigma = 6$  or 200% increase for the 80 years, or at an average velocity of 2.5% a year.

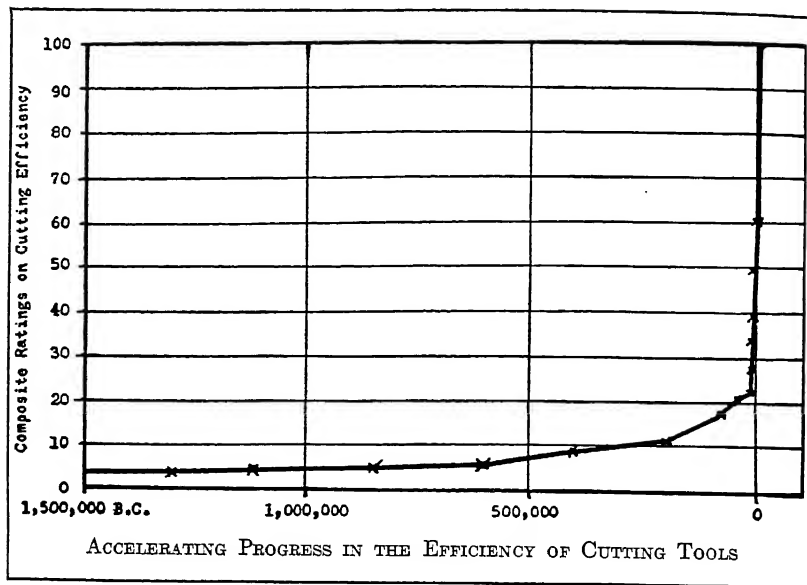
## 3. Competing, $C_p$

"Competition" between prime movers can also be said to exist. Thus of the original animal competitors over the whole period the oxen lose ground, while mules and horses gain ground, but new motor competitors gain ground from scratch. Most of the total gain in horsepower is the process of "progressing,"—farmers now use more horsepower—and only a fraction is competitive in displacing former types of horsepower. Using the percentage form of the index of competing, calculating from readings of the graph, the computation might proceed as follows:

	1850	1930		1930-1850	
	%	Hp. in 10s	%	Change, $+_tV =$	$+_tV^2$
Harvesters	0	2	2.9	+ 2.9	8.41
Trucks	0	22	31.4	+ 31.4	985.96
Electricity	0	2	2.9	+ 2.9	8.41
Gas tractors	0	22	31.4	+ 31.4	985.96
Gas engines	0	3.5	5.0	+ 5.0	25.
Steam engines	0	1.5	2.1	+ 2.1	4.41
Windmills	0	0 +	0	0	0
Horses	62.5%	13	18.6	- 43.9	1927.21
Mules	6.25	4	5.7	- .55	.30
Oxen	31.25	0 +	0	- 31.25	976.56
	100.00%	70	100.00	0	4922.22

$(.5+_tV_{\Sigma}^2)^.5 = 50\% = C_p$  in the 80 years (by Eq. 47, Ch. X). If, with less detail, biological sources and mechanical sources of horsepower are taken as the two competitors,  $C_p = 76\%$ , or an average velocity of competing of just under 1% a year. This means 76% of complete reversal from an initial monopoly of biological sources towards a terminal monopoly of mechanical sources of horsepower used in agriculture.

## S. 26



Ref.: Hart, Hornell, *The Technique of Social Progress*, Henry Holt, 1931, p. 60.

Descriptive formula:  $S_{26} = {}^tT^{-1} : I$

Quantic number = 9;1;0;0

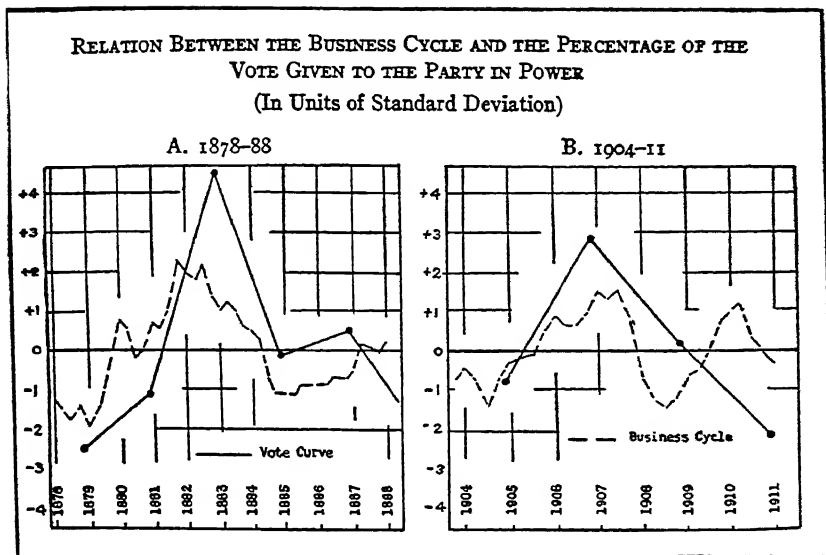
Legend:

$S_{26}$  = The situation  
records for each of

$I$  = an indicator of efficiency of  
cutting tools

${}^tT^{-1}$  = 15 dates from 1,500,000 B.C.  
to date

## S. 27



Ref.: Tibbitts, Clark, "Majority Votes and the Business Cycle," *American Journal Sociology*, Vol. XXXVI, No. 4, January, 1931. Cf. *Social Forces*, Vol. VI, No. 1, Sept., 1927.

Descriptive formula:  $S_{27} = \frac{1}{2} : \frac{1}{2} T^{-1} : \sigma I_1$

Quantic number = 9;1;0;0

Legend:

$S_{27}$  = The situation  
records for each of

$\frac{1}{2}$  = 2 periods

$\frac{1}{2} : \frac{1}{2}$  = (which are not consecutive,  
as their corresponding limiting  
dates are stated)

$\frac{1}{2} T^{-1}$  = subdivided into 18 years

$I_1$  = 2 indicants (of business and  
voting)

$\sigma$  = in sigma units

Comment:

In the tension theory, since the desideratum  $V^0$  = party in power = 1 and the average intensity of desire =  $D$  = one vote,

$P = E$ , i.e., the number of voters = the amount of political tension

The graph studies, but does not explicitly calculate, the correlation between political and economic tensions as measured by these indices. In a business depression economic desiderata (jobs and income),  $V_E$ , decrease (regressing,  $-V_E$ ) and hence tensions increase ( $+E_E$ ), assuming  $P$  and  $D$  unchanged.

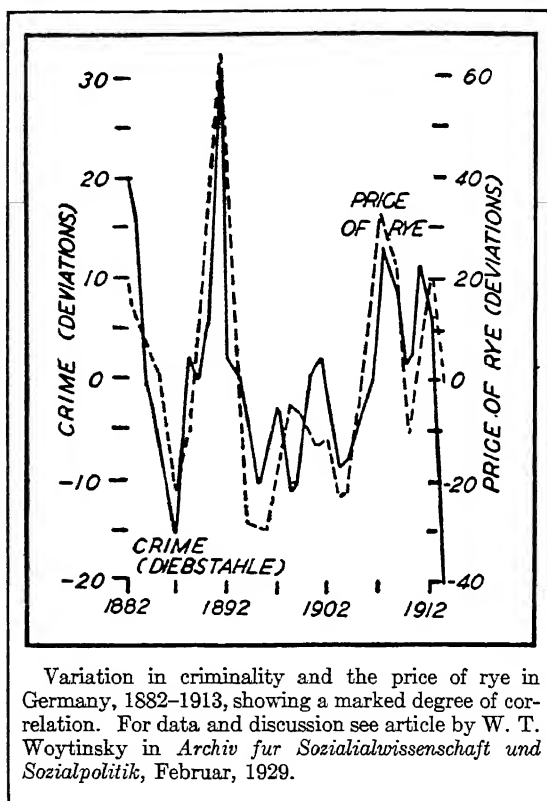
The graph suggests recovering of economic and political tensions asserting the hypothesis that:

$$E_P \cdot E > 0$$

(Eq. 104, Ch. X)

which asserts positive correlation of political and economic attensing as here indirectly measured. The size of  $r$  is not stated.

## S. 28



Ref.: Reinhardt and Davies, *Principles and Methods of Sociology*, Prentice Hall, 1932, p. 588.

Descriptive formula:  $S_{28} = {}^tT^{-1} : I_1$

Quantic number = 9;1;0;0

Legend:

$S_{28}$  = The situation  
records for each of

$I_1$  = 2 indicators  $\left\{ \begin{array}{l} \text{of crime} \\ \text{and rye} \\ \text{prices} \end{array} \right.$

${}^tT^{-1}$  = 30 years

'| = beginning in 1882

Comment:

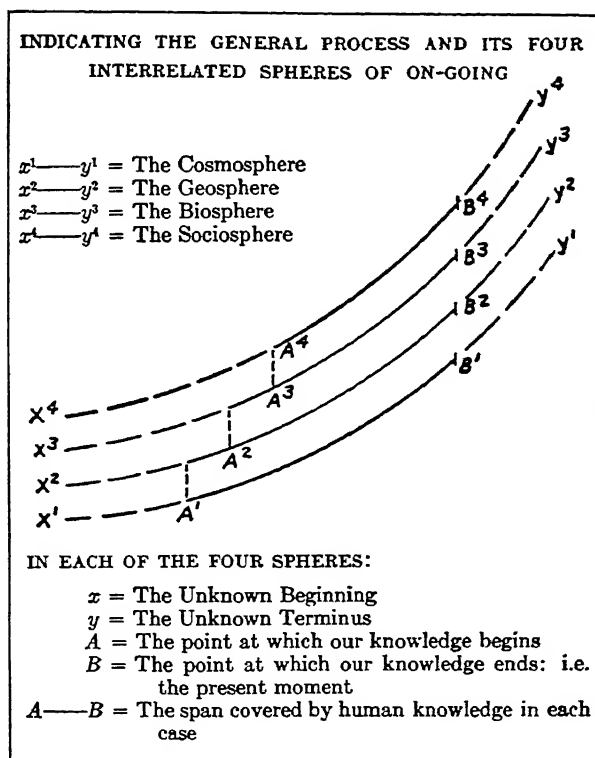
If, in addition to the correlation revealed, there is shown to be a time lag of crime behind price changes (by more accurate methods than mere visual inspection of the graph), then two of the necessary conditions for proving that the economic factors indicated in this situation are partial causes of crime (to the extent of the  $r$ ) are fulfilled.

## S. 28 (Continued)

*Comment on notation:*

The two time curves are merely compared in the graph and hence their indicators are symbolized as aggregated ( $I_1$ ). If the implied correlation were calculated and explicitly stated, the process of recovering Brief-S =  $\tau(I_1)$  would be developed and the two indicators would be symbolized in the standard formula as correlated,  $\tau T^{-1} : I_1 \cdot I_2$ .  $\tau = 9;2;0;0$ .

## S. 29



S. 29 (*Continued*)

*Descriptive formula:*  $S_{29} = \underline{I}_i : {}_t\underline{T}^{-1}$

*Quantic number* = 9;1;0;0

*Legend:*

$S_{29}$  = The situation

in each of

records for each of

${}_t| = 3$  periods

$|_i = 4$  spheres of nature

of

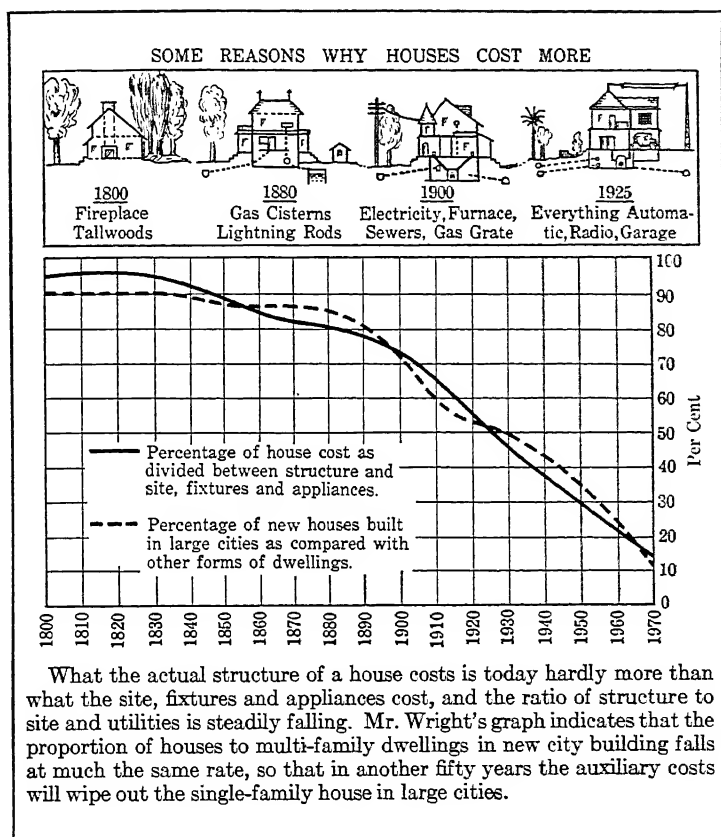
$\underline{I}$  = an implied amount of "on-going"

$\underline{T}^{-1}$  = unspecified duration

*Comment:*

The recorded situation is a second-degree matrix of order  $|_i \times {}_t| = 4 \times 3$ . It has five vectorial dimensions,  $|_i + {}_t| = 4 + 1$ , representing the 4 indicators and time. The vertical ordinate suggests implicitly the amount of change (per period of time along the abscissa). The process illustrated is the most general one of indicating,  $({}_tI)$  (Eq. 28, Ch. X).

## S. 30



Ref.: Wright, Henry, "The Road to Good Houses," *Survey*, Vol. LIV, No. 3, May 1, 1925, p. 166. Original source of illustration: *Journal of the American Institute of Architects*.

Descriptive formula:  $S_{30} = {}^tT^{-1} : {}_{\%}I_1 : I_1^0$

Quantic number = 9;1;0;0

Legend:

$S_{30}$  = The situation  
records for each of

$I_1$  = 2 indicants of housing  
 $\%$  = in percentage terms

${}^t|$  = 17

and

$T^{-1}$  = decades

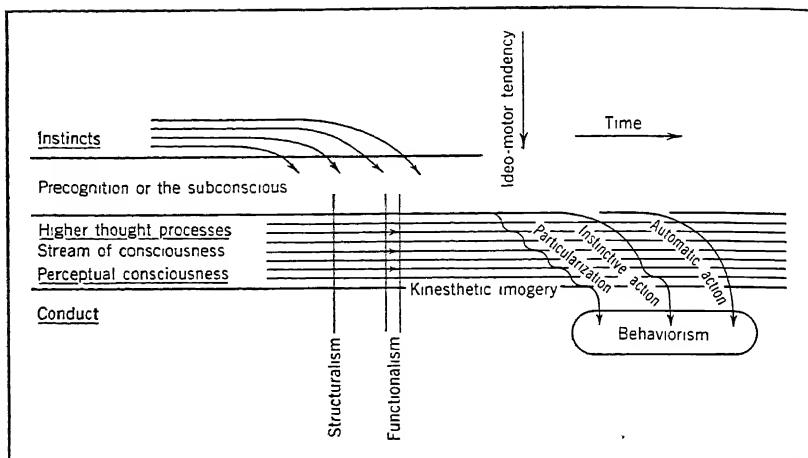
$I_1^0$  = qualitative characteristics of  
the housing

$|$  = beginning in 1800

Comment:

The curve is projected into the future, predicting that, if the current trend continues, by 1970 the cost of the house structure will be only about 15% of the total cost of site, structure, fixtures, and appliances of a home.

## S. 31



Ref.: Thurstone, L. L., *The Nature of Intelligence* Harcourt Brace and Co., 1924, p. 43.

Descriptive formula:  $S_{31} = \text{P} : \text{t}^{\text{t}}\text{T}^{-1} : \text{I}$

Quantic number = 9;1;0;1

Legend:

$S_{31}$  = The situation  
records for

$\text{t}^{\text{t}}\text{I}$  = particular points and one  
period indicated

$\text{P}$  = any person

with a corresponding

$\text{T}^{-1}$  = an unmeasured time period in  
the neural system  
with

$\text{I}$  = ordinal indicant of "ideomotor  
tendency" (varying from the  
subconscious to behavior)

Comment:

The lower limit of societal situations is when the population shrinks to one person. S-theory covers this limiting case and hence can symbolize much of the phenomena of psychology.

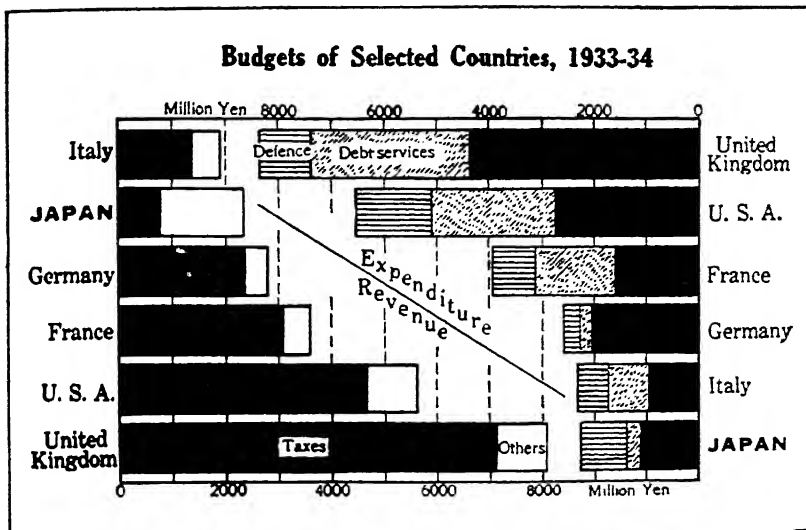
Comment on notation:

The vertical ordinate in the diagram representing "ideomotor tendency" is a continuum without cardinal units but marked by ordinal points ( $\text{I}$ ) identified by such labels as "instincts," "the subconscious," on to "conduct."

The underlining of a capital letter script denotes an indefinite or unidentified individual, i.e., any person taken as a type.

The implied correlation between time and the ideomotor tendency, i.e., the trend for ideas to go from the subconscious into overt action as time goes on, is not developed operationally in the situation as recorded by calculating any form of the correlation coefficient, and hence their symbols are connected by the colon and not the dot. This correlation is  $\text{T} \cdot \text{I}$ .

## S. 32



Ref.: Yano, T. and Shirasaki, K., *Nippon, A Chartered Survey of Japan*, Kukuset-Sha, The First Mutual Building, Kyobashi, Tokyo, 1936, p. 413.

Descriptive formula:  $S_{32} = \underline{P}_p : \iota (IT^{-1})_{i:j}$

Legend:

$S_{32}$  = The situation  
records for each of

$\underline{P}_p$  = 6 nations

$T^{-1}$  = the annual

$I$  = budgets

Quantic number = 9;1;0;1

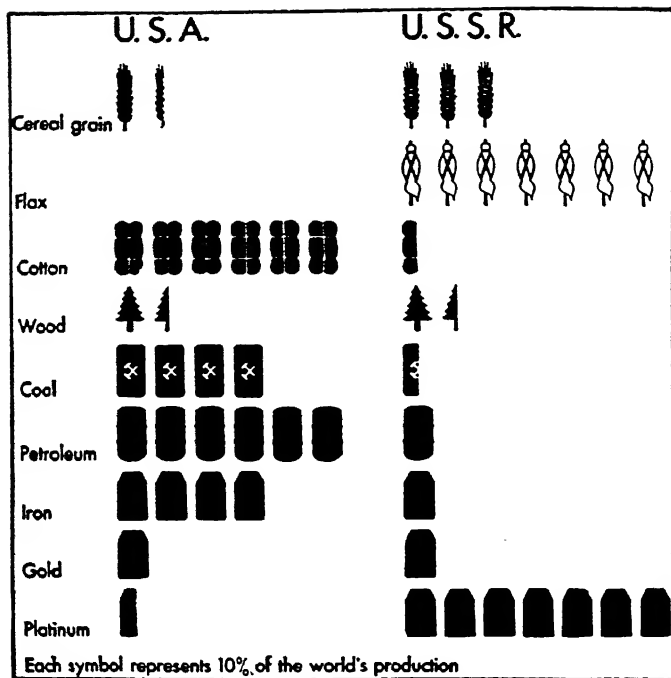
$\iota$  = in 1933-34

classified into

$|_i$  = expenditures and revenues  
and

$|_j$  = their 5 subclasses

## S. 33



Ref.: Duranty, Walter, "U. S. A. and U. S. S. R.," *Survey*, Vol. LXVIII, No. 15, November 1 1932, p. 538.

Descriptive formula:  $S_{33} = \underline{P}_p : \% (IT^{-1})_i$

Quantic number = 9;1;0;1

Legend:

$S_{33}$  = The situation  
records for each of

$\%I$  = production in % units  
of each of

$\underline{P}_p$  = 2 national plurals

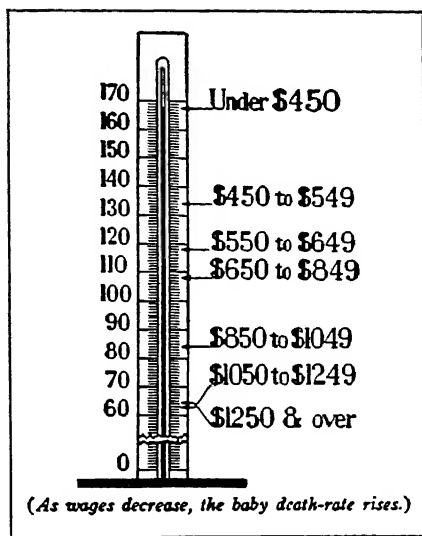
$|_i$  = 9 commodities

$T^{-1}$  = the annual

Comment:

In situations such as this, a coefficient of contingency between nations and commodities would measure the tendency towards regional specialization in production. Insofar as natural resources make for cheaper production in some localities, this is good economically, but is not desired from the political point of view of national self-sufficiency. Change of the coefficient of contingency after a period measures the trend for economic or political considerations to dominate effectively. This coefficient measures for qualitative data not suitable for a correlation coefficient, a variant form of our reco-ordinating process, where  $+r(\%IJ)$  means economic factors are becoming stronger and  $-r(\%IJ)$  means political factors are becoming stronger in the situation as measured. Such measures, refined and qualified more carefully, can be the means of checking hypotheses and predictions on societal issues.

## S. 34



Ref.: Gillin, Dittmer, and Colbert, *Social Problems*,  
Century Co., 1928, p. 317.

Descriptive formula:  $S_{34} = ({}_1I : \%P)T^{-1}$   
Legend:

Quantic number = 9;1;0;1

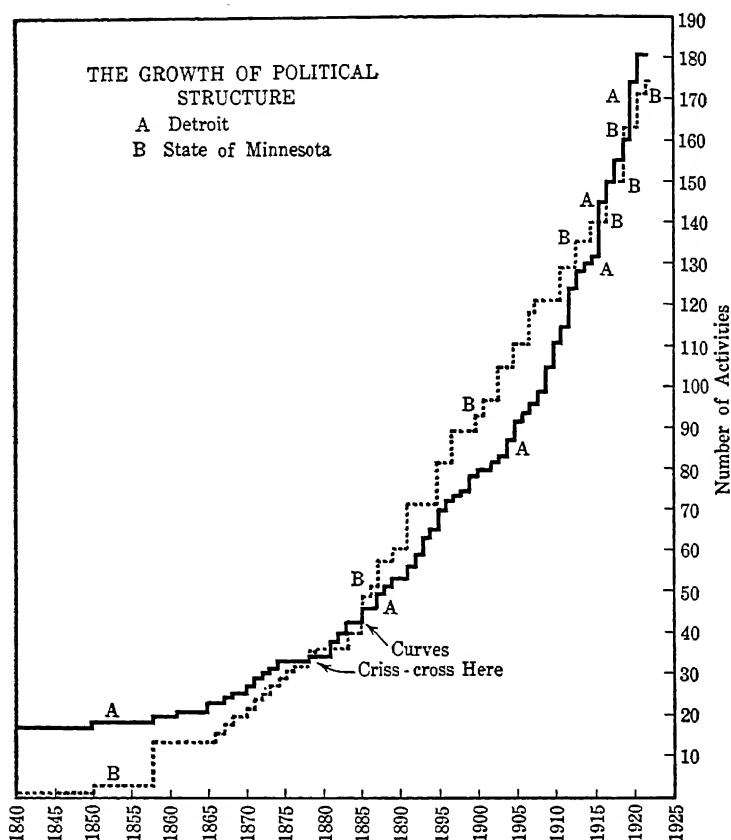
$S_{34}$  = The situation  
records for each of  
 $|$  = 7 class-intervals  
of

$I$  = an indicant of income  
 $T^{-1}$  = per year  
: = the corresponding  
 $\%P$  = infant mortality rate

Comment:

Although operationally undeveloped, since a correlation ratio of income and mortality has not been calculated in these data as presented, this situation could illustrate our reco-ordinating process, the correlation of two dynamic indices.

## S. 35



Ref.: Chapin, Stuart F., "Theory of Synchronous Culture Cycles," *Social Forces*, Vol. III, No. 4, May, 1925, p. 598.

Descriptive formula:  $S_{35} = {}^tT^{-1} : \underline{P}_p : I_{\Sigma 1}^0$

Quantic number = 9;1;0;1

Legend:

$S_{35}$  = The situation  
records for each of

$\underline{P}_p$  = 2 political plurels  
 $|\Sigma 1$  = the number of kinds  
of

${}^tT^{-1}$  = 85 years

'| = beginning in 1840  
for each of

$I^0$  = Government activities

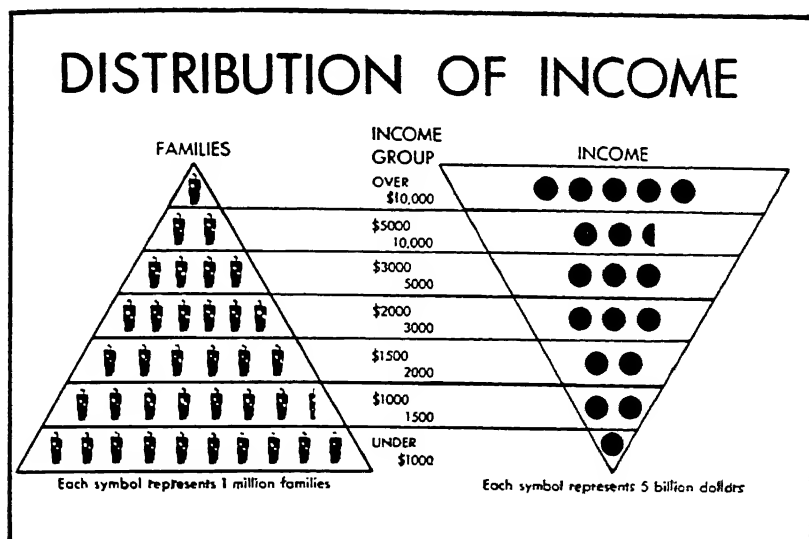
Comment on notation:

The operation of summing, denoted by the  $\Sigma$ , converts, as usual, the aggregative descript into a single number (see S-rule #60 in the Appendix). This could be equivalently written as an indicant where each activity,  $I^0$ , is the unit, so that  $I_{\Sigma 1}^0 = I^{+1}$ . The quantic digit is one 9;1;0;1, as a quantity is denoted in

## S. 35 (Continued)

the indicatory sector and not a mere quality. The descriptive formula, however, shows the quantity to be an increasing number of qualities rather than an increasing amount of some one quantity. This is the process of dissimilarizing, Eq. 26, Ch. X.

## S. 36



Ref.: Pictorial Statistics, Inc., New York.

Descriptive formula:  $S_{36} = {}_i(IT^{-1} : \Sigma_p P)$

Quantic number = 9;1;0;1

Legend:

$S_{36}$  = The situation  
records in

$i|$  = 7 class-intervals of income

$T^{-1}$  = the annual

$I$  = income of families  
:= and the corresponding

$\Sigma_p P$  = number of families

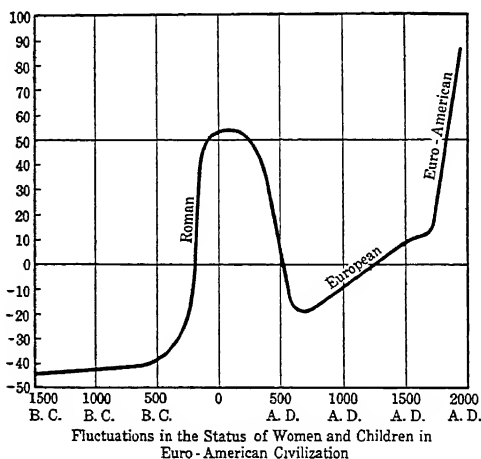
Comment:

For other graphs of the distribution of income see S. 9, Ch. V, the line of equal vs. the curve of unequal distribution; S. 31, Ch. VI, the percentage frequency distribution for each of three educational levels; S. 13, Ch. XI, showing trends for wage income increases to lag behind profits and stock income increases; S. 20, Ch. XII, for the typical skewed frequency distribution curve of incomes in a capitalistic society; S. 40, Ch. X, for wage ogives by regions; and for closely related analyses of costs of living, see S. 32, Ch. XI, and of differential family budgets, S. 47, Ch. X, and of trends of income throughout life for uneducated and educated persons, S. 7, Ch. XII and S. 26, Ch. XII.

## S. 36 (Continued)

Although scattered through five chapters, due to the necessarily linear route of discussing the topics of the quantic classification, these are actually all \* a contiguous region of the quantic solid, S. 33, Ch. II, namely, the  $t;1;0;1$  array of distributions. These are distributions of a population,  $P$ , in a money characteristic,  $I$ , studied with time varying from momentary through changing to accelerating situations (including durations in analyses by age). The exposition proceeding in sequence from topic to topic cannot treat simultaneously all the interweaving relations of societal phenomena. But frequently the quantic solid shows the unity of some section of phenomena which is obscured by the chapter divisions.

## S. 37



Ref.: Hart, Homell, *The Technique of Social Progress*, Henry Holt, 1931, p. 356.

Descriptive formula:  $S_{37} = u, tT^{-1} : \underline{P}, : I$

Quantic number =  $9;1;0;1$

Legend:

$S_{37}$  = The situation

records for each of

$tT^{-1}$  = 7 half millennia

$'$  = beginning in 1500 B.C.

$u$  = including 3 named periods

$\underline{P}$  = the plural for Euro-American women and children

$I$  = an indicant of status

Comment:

To any plural to whom higher status for women and children is a desideratum, the curve records progressing and regressing in the "grading" process (Eq. 33, Ch. X) of the tension theory. If the varying throughout the centuries of this grading were calculated, the process of revarying (or "instability")  $\sigma(tI)$  would be measured (Eq. 58, Ch. X).

\* Except S. 6, Ch. XI ( $\sigma = 8;1;0;0$ ) which is not a distribution of income but an analysis of a causal trend towards increasingly unequal distribution.

## GROUP TYPES AS DIFFERENTIATED BY NATURE OF SOCIAL CONTACT \*

<i>Type of Group</i>	<i>Type of Sensory Contact</i>	<i>Frequency of This Contact</i>	<i>Emotional Intensity of This Contact</i>	<i>Means of Communication</i>	<i>Interdependence and Relationship</i>	<i>Specific Examples of Type</i>
Primary contact group (Intimate contact group)	Face-to-face, direct sense perceptions of auditory, olfactory	Repeated with same persons at many points of contact	Intimate and personal, informal	Oral language, gesture, posture, facial expression	Concrete perceptual elements in the social configuration	Family group, play groups, neighborhood groups
Intermediate group (Superficial contact group)	Face-to-face, direct sensory perception, auditory, visual	Occasional contacts with different persons at different points	Superficial, formal	As above	Same as above, but concrete conceptual elements added	School-room class; audiences; local units of Y.M.C.A., Y.W.C.A., Scouts; church clubs, etc.
Secondary contact group (Artificial contact group)	Derivative or indirect sense perception, mediated by mechanical means	Contacts at infrequent intervals	Highly impersonal through artificial devices of communication	Telephone, telegraph, radio, printed materials	Abstract conceptual, or symbolic elements in the social configuration	Headquarters of a national society, regional council, board of directors, executive committee, etc.

\* The studies of small groups by using a measure of group acceptance as an indicator of group cohesion have been advanced by the work of Newstetter. Consult Newstetter, W. I., and Feldstein, M. J., *A Brief Summary of the Waukegan Camp Research Project, 1931-1932* (mimeographed), Western Reserve University, 1932, p. 30; Chapin, F. Stuart, *Cultural Change*, pp. 308-309.

Ref.: Chapin, F. Stuart, *Contemporary American Institutions*, Harper and Brothers, 1935, p. 102.

S. 38 (*Continued*)

Descriptive formula:  $S_{38} = \underline{P}_p : q : ({}^{\Sigma}T^0, \underline{I}_i)$

Quantic number = 9;1;0;1

Legend:

$S_{38}$  = The situation

${}^{\Sigma}T^0$  = the frequencies of contact

records for each of

and

$\underline{P}_p$  = 3 groups  $\left\{ \begin{array}{l} \text{primary} \\ \text{intermediate} \\ \text{secondary} \end{array} \right.$

$\underline{I}_i$  = 4 sets of ordinal indicants

$|_q$  = with corresponding sub-  
plurels ("Specific Examples  
of Type")

$\left\{ \begin{array}{l} \text{Type of contact} \\ \text{Emotional intensity} \\ \text{Means of communication} \\ \text{Interdependence} \end{array} \right.$

*Comment on notation:*

The "frequency of contact" is an unspecified *number* of dates not necessarily in sequence, which is denoted in the descriptive formula by  ${}^{\Sigma}T^0$  but in the quantic by  $T^{-1}$ , since a plural date implies something going on in time and not a mere instant. The operation of counting the dates or summing the points in the time dimension converts a date into a period, a point into a line. This number of dates (if more definitely stated) would measure the periodizing process, the number of dates in a period.

Note the indicatory case script denoting a point which, combined with the underlined indicant denoting indefiniteness, denotes an ordinal quantity, since in ordinals the points on a scale are specified but with unknown intervals between them.

## S. 39

That an average per capita incidence of taxation of LP.1.350 is low, while an average of LP.4.270 is high, becomes evident from a comparison with the per capita figures for other countries, as follows:

	LP.
India (1927-28)	0.173
Bulgaria (1927-28)	1.088
Soviet Russia (1929)	1.200
Palestine (non-Jews) (1930)	1.350
Turkey (1929)	1.400
Japan (1929-30)	1.636
Egypt (1929)	1.700
Poland (1929-30)	1.895
Yugoslavia (1929)	1.900
Union of South Africa (1927-28)	2.700
Hungary (1929)	3.100
Switzerland (1929)	3.137
Finland (1930)	3.217
Italy (1929-30)	3.423
Czechoslovakia (1930)	3.585
Palestine (Jews) (1930)	4.270
Spain (1930)	4.464
Chile (1929)	4.500
Sweden (1929-30)	5.132
France (1929)	5.300
Denmark (1929-30)	5.484
Norway (1930-31)	5.567
Netherlands (1929)	5.872
Austria (1930)	5.873
Belgium (1930)	5.930
United States of America (1928-29)	6.090
Canada (1927-28)	7.780
Australia (1927-28)	9.100
Great Britain (1928-29)	15.000

Ref.: Granovsky, A., *The Fiscal System of Palestine*, "Mishar ve-Taasia" Publishing Co., Ltd., Jerusalem, 1935, pp. 297-298.

*Descriptive formula:*  $S_{39} = (IP^{-1}T^{-1})_p : t'$

*Quantic number* = 9;1;0;1

*Legend:*

$S_{39}$  = The situation  
records

$|_p$  = 29 nations

for

I = the taxation

$| : t' =$  a corresponding period for  
each

$P^{-1}$  = per capita

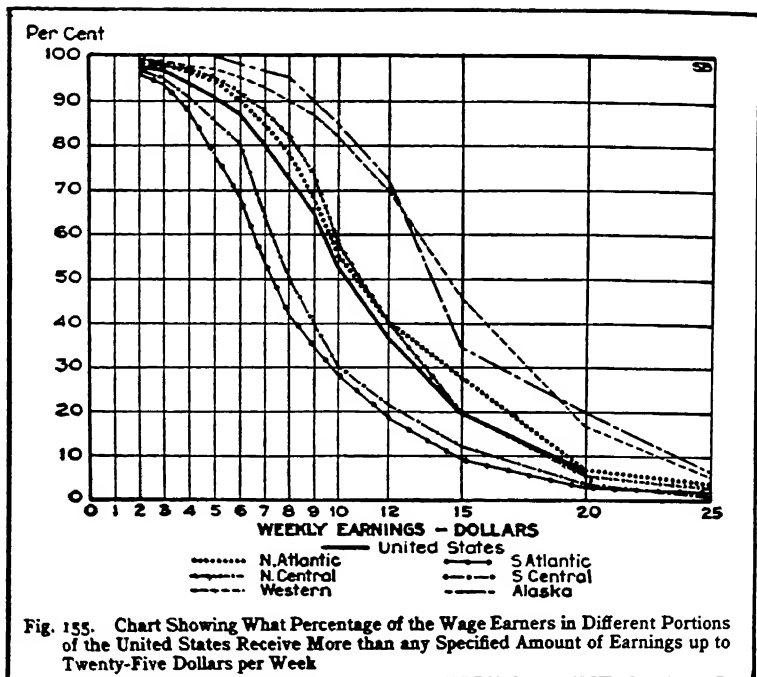
$T^{-1}$  = per annum

of

*Comment:*

The mean per capita incidence of taxation for these 29 countries is 4.20 £P (= \$20.41 in 1930) and measures the indicating process  $(_1I)$  (Eq. 28, Ch. X), while the standard deviation is 2.95 £P (= \$14.33) which measures the reordinating process  $^{\circ}(_1I)$  (Eq. 46, Ch. X) in national taxation.

## S. 40



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 181.

Descriptive formula:  $S_{40} = {}_i(IT^{-1}) : \%P_{p,\Sigma p}$

Quantic number = 9;1;0;1

Legend:

$S_{40}$  = The situation  
records for each of

$\%P$  = percentage frequency of persons

$|_i|$  = 23 class-intervals  
of

in each of

$I$  = earnings (\$)

$|_p|$  = 6 U.S. regional plurels

and in

$T^{-1}$  = per week

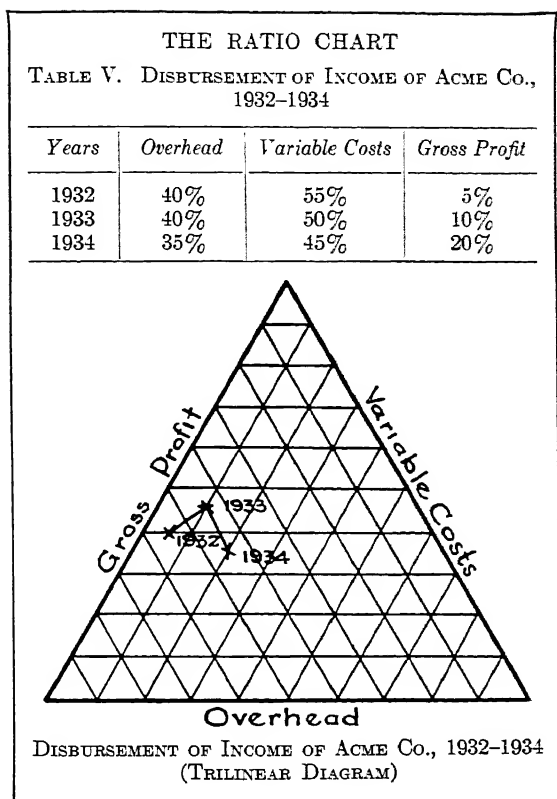
$|\Sigma p|$  = the U.S. as a whole

$:$  = the corresponding

Comment:

The tendency of wages to vary regionally from highest in the west to lowest in the south is not developed, in the situation as recorded, to the point of calculating the correlation.

## S. 41



Ref.: Arlsin and Cotton, *Graphs: How to Make and Use Them*,  
Harpers, 1936, p. 89.

Descriptive formula:  $S_{41} = \underline{P}_t : {}_tT^{-1} : \%I_1$   
Legend:

Quantic number = 9;1;0;1

$S_{41}$  = The situation  
records

'| = beginning in 1932

$I_1$  = 3 income indicants

$\underline{P}_t$  = for the Acme Co.  
for each of

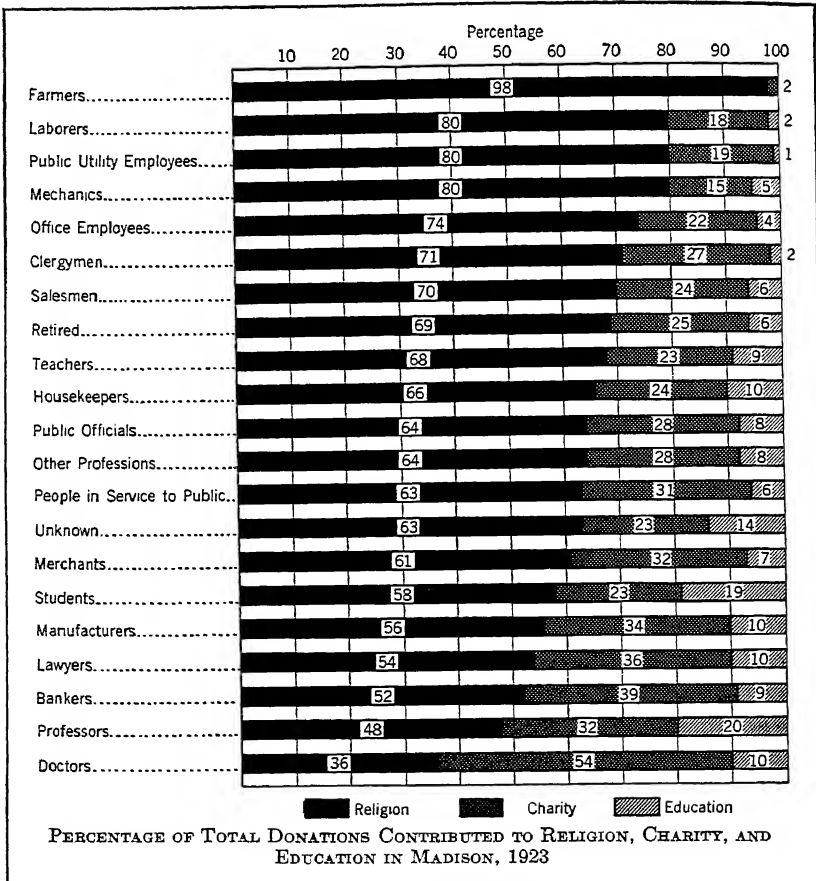
$\%|$  = in percentage units

${}_tT^{-1}$  = 3 years

Comment:

This unusual type of trilinear diagram is difficult to interpret. It is included here as further evidence of the comprehensiveness of S-notation in symbolizing any and all quantitative sets of data in the social sciences, regardless of their form of presentation in graph, tabulation, or prose.

## S. 42



Ref.: Gillin, John, "Public Welfare and Social Work," *Social Forces*, Vol. X, No. 3, March, 1932, p. 367.

Descriptive formula:  $S_{42} = P_p : \% (IT^{-1})_i$

Quantic number = 9;1;0;1

Legend:

$S_{42}$  = The situation

I = donations

records for each of

to each of

$P_p$  = 21 occupational plurels in  
Madison

$|_i$  = 3 "causes"

$\%$  = in % units

$T^{-1}$  = the annual

Comment:

From the contributor's point of view the percentage ( $\%D$ ) given to each of the three types of charity measures that plurel's relative valuation of each of the three. This is a doubly aggregated evaluating process—the evaluating is aggre-

## (S. 42 Continued)

gated once for the three values ( $V^0 = 3$  charities) and again for the 21 plurels ( $|_p$ ). The tension theory equation here is:

$$(\frac{P^0}{\%}D = V^0E)_{p:v} \quad (\text{Eq. 105a, Ch. X})$$

$$(\%D = E)_{p:v} \quad (\text{Eq. 105b, Ch. X})$$

Since the desiderata are each unitary, "a charity," the tension of a plurel towards each desideratum is identical with the intensity of desire of that plurel for that desideratum.

As usual, at the present primitive stage of development of societal units, tension, E, is most easily measured in those situations where one or more of the other three factors are unity, thus simplifying the problem of compounding diverse units.

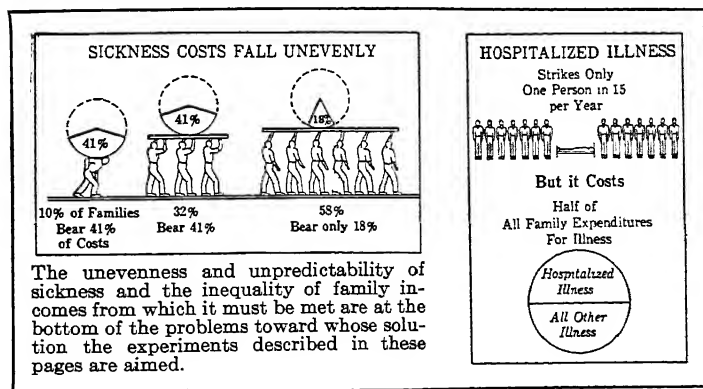
From the point of view of the contributee, the donations are the desideratum, V, desired, while D is unmeasured and  $|_p =$  the 3 charity plurels. Defining the population as the contributees, results in the relative contributions measuring the process of competing. Cp, the index of effective competing (Eq. 47a, Ch. X) varies from 56% among the farmers, where religion has almost a monopoly of their donations, to a Cp of 14% among professors, where the three charities compete somewhat more equally.

For Professors (by Eq. 50, Ch. X)

	V'	100/	$\pm\%V$	$\pm\%V^2$
Religion	48	-33.33	+ 14.67	215.21
Charity	32	-33.33	- 1.33	1.77
Education	20	-33.33	-13.33	177.69
				<u>394.67</u>

$$(.5 \times 394.67)^.5 = 14$$

## S. 43



Ref.: Davis, Michael M., "Change Comes to the Doctor," *Survey Graphic*, Vol. XXIII, No. 4, April, 1934, p. 165. Charts from *A Picture-Book about the Costs of Medical Care*, Julius Rosenwald Fund, Chicago.

S. 43 (*Continued*)

*Descriptive formula:*  $S_{43A} = \%P_p : \% (IT^{-1})$

*Quantic number* = 9;1;0;1

*Legend:*

$S_{43A}$  = The situation  
records for each of

: = the corresponding

$T^{-1}$  = dynamic (annual?)

$P_p$  = 3 plurels of families

I = indicants of cost

$\%|$  = stated in percentage units

$\%|$  = in % units

*Descriptive formula:*  $S_{43B} = T^{-1} : \%P : \%I$

*Quantic number* = 9;1;0;1

*Legend:*

$S_{43B}$  = The situation  
records for

$\%P$  = % of persons hospitalized

$T^{-1}$  = the annual

$\%I$  = a % of family expenditures

*Comment on notation:*

Although these two graphs have the same quantic number, they are two S-situations and not one, because they cannot both be unambiguously described by a single descriptive formula.

## S. 44

## STANDARD DEVIATIONS

Test	Feeble Minded	College Juniors		Orphan 6th Grade			12 Yr. Olds	2nd-6th Grade	Hindus
		Short Time	Unltd Time	1st Trial	2nd Trial Reg. Time	2nd Trial Unltd Time			
a	b	c	d	e	f	g	h	i	j
IB. Easy learning	9.68	8.77		8.16	8.96		9.89	8.94	7.41
IC. Hard learning	7.	7.74		7.48	9.57		8.09	8.51	6.724
2. Matching	2.70	2.01	1.69	2.57	2.15	2.27	2.69	3.02	2.584
3. Dots	3.91	2.55	2.12	3.12	3.42	3.42	2.79	3.14	3.942
4. Rhythms	5.59	8.21	9.18	5.36	7.48	8.26	5.26	5.68	5.797
5. Mazes	1.35	2.61	1.72	3.3	4.04	4.32	3.39	3.2	2.729
6. Cubes	4.6	14.16	6.4	5.6	6.96	7.04	6.19	6.0	4.864
7. Similarities	7.31	15.09	12.	9.6	10.84	10.8	9.62	8.92	7.432
8. Faces	5.32	8.2	4.55	12.69	12.	11.96	11.64	11.70	6.888
9. Time series	5.71	7.27	5.	6.54	7.32	7.46	7.29	7.15	5.260
10. Analogies	4.04	6.12	5.8	4.94	5.62	5.62	4.61	4.54	4.098
11. Geometric series	12.3	8.72	5.8	10.67	8.88	16.88	12.02	13.09	11.667
Scale	72.	68.85	52.36†	50.25	58.65	59.81	66.23	68.98	*
M.A. (months)	18.78			14.48			14.79	19.78	34.**
C.A. (months)				13.68			3.73	19.62	18.82
E.A. or "T." or "P."		9.84§		13.64			15.12	19.60	
School grade							.869	1.173	1.644
Length of residence								3.881 mos.	
E.Q.								3.06	

\* Not comparable with the other sigmas as the tests were combined with sigma weights.

† "Regular" time sigma is 61.07.

§ This is in Princeton sigma units, where  $\sigma = 10$  points. It is the  $\sigma$  of the higher of the Princeton or Thorndike intelligence tests for college admission.

\*\* Punjabi Binet Scale.

Ref.: Dodd, Stuart C., *International Group Mental Tests*, Princeton Univ. Store, 1926, p. 40.

Descriptive formula:  $S_{44} = \underline{P}_D : {}^tT^{-1} : {}^eI_i$       Quantic number = 9;1;0;1

Legend:

$S_{44}$  = The situation

records for each of

$\underline{P}_D$  = 6 plurels

and for each of

${}^tT^{-1}$  = several occasions and periods

${}^eI$  = standard deviations

of

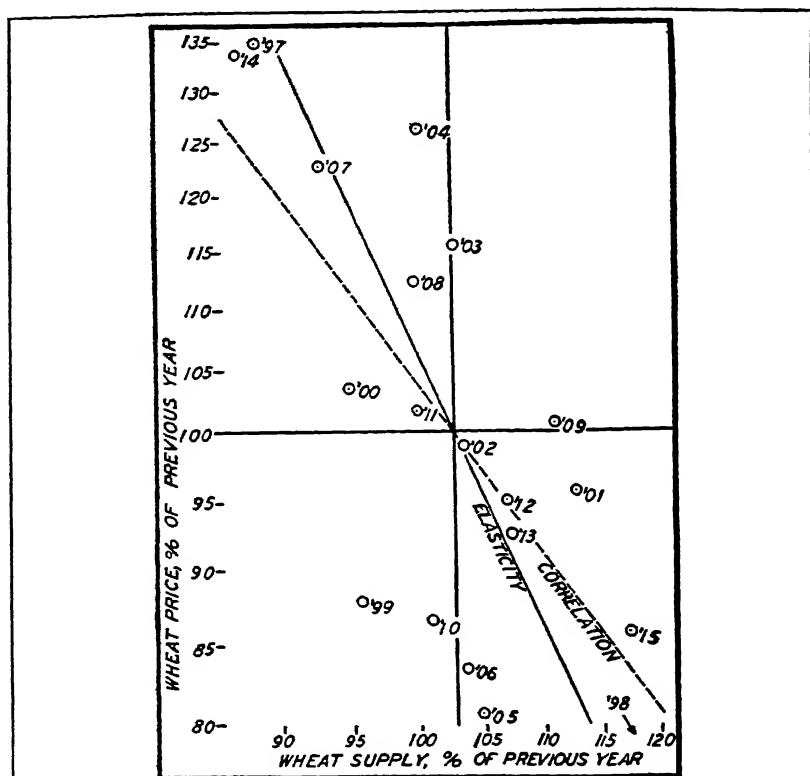
$|_i$  = 19 indicants relative to intelligence

## S. 44 (Continued)

## Comment:

The process of redispersing is illustrated by the differences between columns e and f, the sigmas on first and second trials. Also, as the score is a desideratum, desired by the testee and achieved by a period of activity, the sigma of scores could measure the process of competing.

## S. 45



The law of supply and demand: price of wheat (December 1) as influenced by supply (crop and carry-over), 1897-1915. Based on percentages of given year to previous year (link relatives of price and supply), showing that in general a supply increase is followed by a price decrease, and vice versa. Price variability is greater than supply variability; that is, the market is inelastic ( $\eta = .45$ ). Correlation,  $r = -.74$ ; slope of elasticity regression line (fitted by least squares direct to log points),  $b = -2.210$ . The figures for 1898 were extreme (supply, 131%, price, 71%) and do not appear on the chart.

## S. 45 (Continued)

Descriptive formula:  $S_{45} = {}_iI :: {}_iI : {}^tT^0$ 

Quantic number = 9;2;0;0

Legend:

 $S_{45}$  = The situation ${}_iI$  = a demand indicant

records for every value of

the corresponding

 ${}_iI$  = a supply indicant ${}^tT^0$  = frequency of instances, or  
dates

:: = cross-classified with

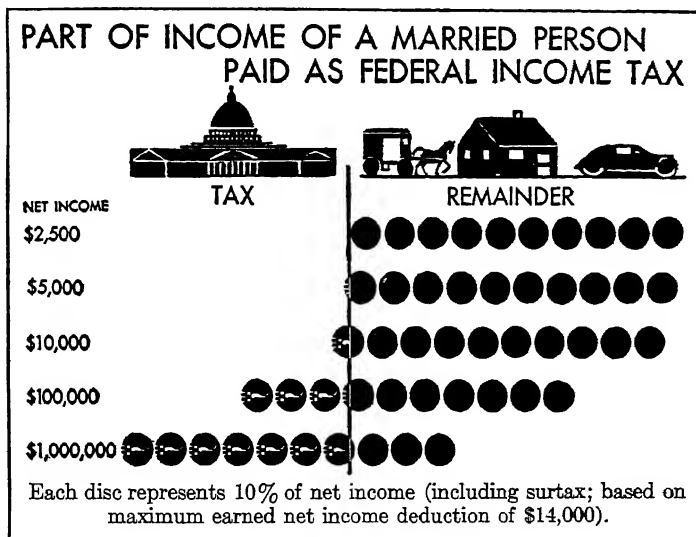
## Comment:

Note the unusual unit of frequency, a date. The calculative formula is:

$$r = t^{-1} \sum_{i=1}^t I_i \sigma I_j = -.74 \quad (\text{Eq. 106, Ch. X})$$

The quantic is  $T^{-1}$ , since the dates are in sequence. It is a case of recovarying, a time series correlation, in scattergram form.

## S. 46



## S. 46 (Continued)

Descriptive formula:  $S_{46} = {}^{\perp}P : ({}_iI :: \%I_i) \underline{T}^{-1}$ 

Quantic number = 9;2;0;1

Legend:

 $S_{46}$  = The situation

records for any

 $I_i$  = 2 types of  
expenditure{ taxes  
and  
other ${}^{\perp}P$  = typical married person

stated in

 ${}_iI$  = the size of income in one of 5  
classes $\%$  = percentage units

cross-classified with

 $\underline{T}^{-1}$  = per year

## S. 47

EXPENDITURES OF FAMILIES IN VARIOUS INCOME GROUPS BY  
PERCENT

Income Class.....	\$1,000- 1,499	\$2,000- 2,499	\$4,000- 4,999	Total
No. of Fams.....	144	266	49	1,038
<i>Expenditures</i>				
Food.....	26.5	22.3	15.5	21.4
Clothing.....	7.3	7.2	8.1	7.6
Auto maintenance and purchase...	15.1	12.8	11.5	12.6
Medical and dental care.....	4.4	3.0	2.8	2.9
Other "living costs".....	21.5	22.0	22.7	21.4
TOTAL "living costs".....	74.8	67.3	60.7	65.9
Benevolences.....	6.7	7.7	8.9	8.0
Vacation, travel, amuse.....	2.2	2.9	3.7	3.2
Books, mag., tuition.....	4.0	4.4	5.0	4.9
Other "advancement costs".....	2.1	3.2	2.9	2.8
TOTAL "advance. costs".....	15.0	18.2	20.5	18.9
Life insurance.....	6.3	7.1	8.6	7.2
Savings.....	3.1	1.2	8.0	6.5
Other "investment".....	0.8	1.2	2.2	1.5
TOTAL "investment".....	10.2	14.5	18.8	15.2
TOTAL Expenditures.....	100.0	100.0	100.0	100.0

Ref.: Leiffer, Murray H., "Income and Standards of Living in the Ministry," *Sociology and Social Research*, Vol. XVII, No. 5, May, 1933, p. 451.

S. 47 (*Continued*)

*Descriptive formula:*  $S_{17} = {}_s(IT^{-1})_{1,\Sigma 1} : ({}_sP, : {}_sI_1 : k)$  *Quantic number* = 9;2;0;1  
*Legend:*

$S_{17}$ = The situation	:: = and cross-classified
records each of	with
$ _{1,\Sigma 1}$ = 4 classes (and their composite)	$ _j$ = 3 classes of expenditures
of	and
$T^{-1}$ = annual	$ _k$ = 12 subclasses
${}_sI$ = income in dollar units	${}_s $ = in percentage units
with their corresponding	
${}_sP$ = number of family plurels	

*Comment:*

The tendency of proportions of family budgets spent for living costs vs. luxuries ("Engel's law") to vary with size of income is summarized, for one measure, in a coefficient of contingency of .78. The situation is a variation of our reco-ordinating process, as it is a correlation between 2 dynamic indicants.

*Comment on notation:*

Note the use of the class-interval script to specify the units, the use of the summation sign,  $\Sigma$ , to denote the sum of the classes of the class script preceding it; and the repeated colon before the second I, which together with the first colon outside the parenthesis converts subclassification into cross-classification (:,:), which is partially developed correlation.



7. Similarities	.45	.65	.80	.72	.90	.66	.92	.88	.95	.80	.96	.93
8. Faces	.67	.63	.85	.72	.89	.84	.94	.88	.94	.91	.98	.93
9. Time series	.31	.63	.73	.71	.89	.49	.88	.87	.94	.66	.94	.96
10. Analogies	.50		.78	.61	.70	.69	.91	.80	.82	.82	.95	.89
11. Geometric series	.53	.42	.72	.81	.87	.72	.90	.92	.93	.84	.93	.96
Scale	.58	.81	.83	.88	.93	.78	.93	.95	.96	.88	.97	.98
Scale **		.90	.91	.94	.96		.97	.98				

\* Undistributed. Many perfect scores in unlimited time.

† For the observed sigma of 14.48 Binet mental months, the "standard" sigma of 25 months was substituted by formula # 186, Kelley T. L. Statistical method.

‡  $R = \frac{2r}{1+r}$   $r =$  "standard" range  $r$ .

\*\* Corrected by the Spearman Brown formula (3).

Ref.: Dodd, Stuart C., *International Group Mental Tests*, Princeton Univ. Store, 1926, p. 42.

S. 48 (*Continued*)

*Descriptive formula:*  $S_{48} = {}^tT^{-1} : \underline{P}_p : (I_{i..})_{i:k}$       *Quantic number* = 9;2;0;1

*Legend:*

$S_{48}$  = The situation

$|_i = 12$  tests

records on

$(I)_{i..}$  = reliability coefficients

${}^tT^{-1} = 2$  trials

in

and under

$|_j = 3$  classifications as to range

${}_t| = 2$  time limits

and

for

$|_k = 2$  subclassifications as to

$\underline{P}_p = 2$  plurels

scoring  $\left\{ \begin{array}{l} \text{rows} \\ \text{rotators} \end{array} \right.$

for each of

*Comment:*

The recorrelating processes of (1) reranking is measured by the repetition reliabilities columns b, g, and k; (2) reco-ordinating is measured in the other columns when the score is considered as a dynamic indicant of points earned per test period; and (3) recodispersing is measured by the differences between columns d and e, or h and i, or l and m, which show change in correlations.

SUMMARY OF INTER-AREA MOVEMENTS, 1925-1936  
1922, 1923, 1924 Policies

Area of Origin	Area of Location																			
	A	B	C	D	E	F	G	H	J	K	L	M	N	P	W	R	S	T	V	Y
A	645	61	16	0	6	2	6	16	7	6	1	0	6	6	54	58	4	6	8	6
B	63	242	17	0	3	1	3	8	9	0	1	3	3	12	25	0	1	1	0	1
C	19	15	154	6	11	0	3	5	1	1	3	5	1	7	6	1	1	4	0	0
D	10	5	6	33	5	1	1	0	0	1	1	1	6	21	1	3	0	1	1	1
E	6	4	13	3	151	0	2	1	0	0	1	0	0	0	0	0	0	0	0	0
F	0	1	0	1	1	3	47	2	0	0	0	0	0	0	7	4	0	0	0	0
G	4	9	3	2	1	1	3	74	1	1	0	0	2	8	12	1	1	0	0	0
H	21	5	5	0	0	0	2	0	33	0	0	1	4	3	4	0	0	2	0	0
J	9	1	0	4	0	0	0	1	2	43	6	0	2	9	3	4	2	2	0	0
K	9	1	0	3	1	0	0	1	1	6	37	7	1	9	9	3	2	2	0	0
L	3	6	0	3	1	0	0	1	1	0	9	42	8	5	2	7	2	4	0	0
M	3	4	0	1	2	0	0	2	5	2	1	3	67	6	5	2	1	6	0	1
N	7	3	0	1	5	0	3	0	0	6	2	2	5	443	7	13	1	6	2	2
P	16	3	2	1	8	1	1	0	0	5	4	2	8	28	448	10	7	9	4	3
W	52	12	2	0	6	0	1	8	1	5	4	2	4	20	17	341	19	6	3	0
R	15	1	7	0	1	0	0	0	0	2	7	4	4	7	5	5	45	3	3	2
S	5	1	1	1	0	0	0	0	0	1	1	0	1	4	3	1	0	80	7	3
T	4	2	1	0	1	0	0	0	0	0	0	0	0	2	2	2	0	1	40	2
V	2	1	1	1	1	0	0	0	0	0	0	0	0	2	3	0	1	2	1	1
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	0	1	2	1	2

Ref.: Moore, Elton II., "Mobility of Insurance Policy Holders," *American Sociological Review*, Vol. III, No. 1, Feb., 1938, p. 71.

## S. 49 (Continued)

Descriptive formula:  $S_{49} = (P_p :: P_p)T^{-1}$ 

Quantic number = 9;0;0;2

Legend:

 $S_{49}$  = The situation  
records $|_p$  = 20 regional plurels

cross-classified with

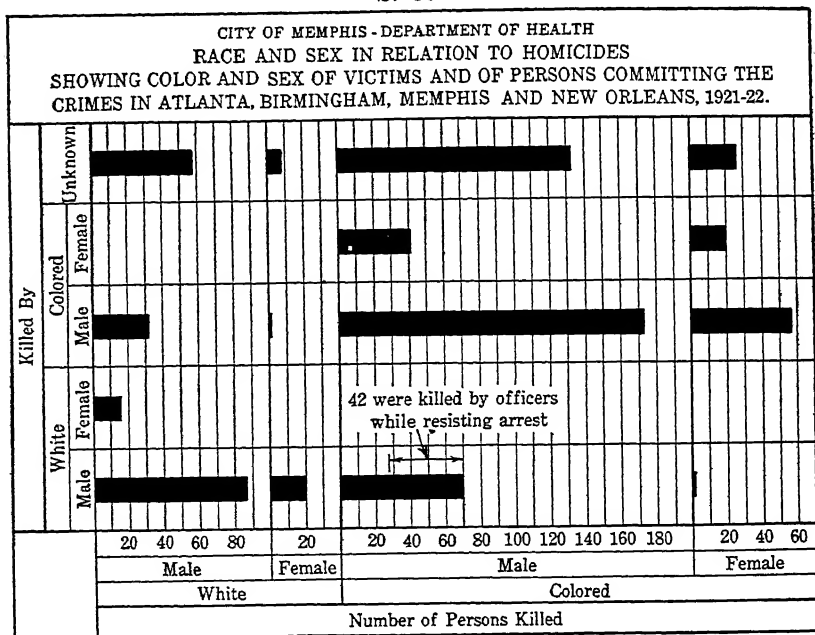
 $T^{-1}$  = the decennial $P_p$  = themselves as immigrants to  
those 20 regions $P$  = emigrants

from

Comment:

The full "migration matrix," a subform of the populational interaction matrix, is shown. This exact analysis provides the detailed "structure" of the migrating process in this situation.

## S. 50



Ref.: Durrett, J. D. and Stromquest, W. G., "Preventing Violent Death," *Survey*, Vol. LIV, No. 8, July 15, 1925, p. 436.

Descriptive formula:  $S_{50} = \underline{P}_{p'} : r : s : t'(PT^{-1})_{p'} : R : s$ 

Quantic number = 9;0;0;2

Legend:

 $S_{50}$  = The situation  
records $|_s$  = subdivided into 2 sex plurels  
with their corresponding $\underline{P}_{p'}$  = the murderers of 4 cities $P_{p'}$  = number of persons murdered $|_r$  = subdivided into 2 racial  
plurels $t'T^{-1}$  = in 1921-22 $|_{R:s}$  = subdivided by race and sex

S. 50 (*Continued*)*Comment:*

The graph shows the full matrix of individual conflicting (Eq. 10, Ch. X) as conditioned by the two racial and sex characteristics. It illustrates a subform of the depopulating process which is operationally fully developed in that both parties of the interacting are recorded in a cross-classification.

*Comment on notation:*

Note the use of the capital letter descript to denote the recipient of interaction, while the small letter descript denotes the expresser or doer of the action.

## S. 51

## CHILDREN PER MARRIAGE IN PRUSSIA, 1875-90, ACCORDING TO RELIGION OF CONTRACTING PARTIES

Creed of Fathers	Creed of Mothers		
	Evangelical	Catholic	Jewish
Evangelical . . . . .	4.35	3.30	1.78
Catholic . . . . .	3.34	5.24	1.66
Jewish . . . . .	1.58	1.38	4.21

*Ref.:* Holmes, Samuel J., *The Trend of the Race*, Harcourt, Brace and Co., 1921, p. 357.

*Descriptive formula:*  $S_{51} = \underline{P}_p :: \iota(P, \underline{P}_{\iota}^{-1} T^{-1})_p$       *Quantic number* = 9;0;0;29  
*Legend:*

$S_{51}$  = The situation  
 records for each of

$|_p$  = 3 religious plurels  
 of

$\underline{P}$  = Prussian fathers  
 $::$  = cross-classified with

$\underline{P}_{\iota}^{-1}$  = mothers  
 $P_{\iota}$  = the number of children (per  
 mother)

for

$\iota T^{-1}$  = the period 1875-90

*Comment:*

The adpopulating process, developed for religious interrelations of parents, is illustrated.

*Comment on notation:*

This situation presents an interesting case of the issue between assigning a quantic of  $P^{+2}$  or of  $P^{+3}$ . There is no intrinsic objection to  $P^{+3}$ , and it would identify the populational interrelation matrix, overlapping plurels, mobility, etc.

In the case of overlapping membership, it can be argued that the members are in two (cross) classifications, hence two vectors, and so  $P^{+2}$ . But in S. 51 above the children are neither fathers nor mothers. 3 different plurels are involved, not one. To call the children doubly classified by two indicals ("mother's religion" and "father's religion") is possible, making a contingency table of it. But interrelations of parents are clearly involved, making this less accurate than  $P^{+2}$ . The ratio adds a digit of 9 for the  $P^{-1}$ . Since the situation can be represented, without adding a new rule, by a quantic of 29 by the regular rules, this is preferred on grounds of parsimony.

## S. 52

TABLE OF ATTITUDES 1

<i>Towards Elsa:</i>	<i>Sympathy</i>	<i>Anger</i>	<i>Fear</i>	<i>Dominance</i>
Maude.....	4	0	0	0
Gladys.....	2	1	0	1
Joan.....	1	2	0	1
Virginia.....	0	2	0	2

The attitudes taken by Elsa towards the four individuals is tabulated in the accompanying Table of Attitudes 2:

TABLE OF ATTITUDES 2

<i>Elsa towards:</i>	<i>Sympathy</i>	<i>Anger</i>	<i>Fear</i>	<i>Dominance</i>
Maud.....	4	0	0	0
Gladys.....	2	1	1	0
Joan.....	2	1	1	0
Virginia.....	0	4	0	0

Ref.: Moreno, J. L., *Who Shall Survive?* Nervous and Mental Disease Publishing Co., 1934, p. 183.

Descriptive formula:  $S_{52} = 'P :: {}^pP : I_1^0 : {}^tT^{-1}$       Quantic number = 9;0;0;2  
 Legend:

$S_{52}$  = The situation

records

'P = a reform schoolgirl, Elsa,  
 interacting with

${}^pP$  = 4 other girls

by expressing

$I_1^0$  = 4 kinds of attitude

{ sympathy  
 fear  
 anger  
 dominance  
 each upon

${}^tT^{-1}$  = a stated number of occasions  
 at unstated intervals

Comment:

The table entries showing the number of dates (occurrences) of each girl's expressing a given type of emotion towards each other girl illustrates the periodizing process fully developed in an interacting matrix.

Note that  ${}^tT^{+1}$  would connote occasions-in-sequence, as T, the duration from some common origin date, is stated. T is  $|^{-1}$  to show that the  $I_1^0$  are dynamic.

## S. 53



## MOVIES

THE movies have become a social-psychological problem because of their influence on children. While most adults are able to withstand the attack on the emotions, children, easily excited, may develop permanent neurotic conditions. But movies affect adults in their standards of behavior.



## S. 53 (Continued)

Descriptive formula:  $S_{53} = {}^pP_p : {}^qP : (I^0T^{-1})_i$ 

Quantic number = 9;0;0;2

Legend:

 $S_{53}$  = The situation  
records $I^0T^{-1}$  = stimulating (i.e., qualitatively acting upon) ${}^pP$  = 20 persons $|_i$  = in an indefinite variety of ways

and

 ${}^qP$  = 15 of those persons $|_p$  = 7 groups

Comment:

The situation illustrates two important uses of S-symbols to represent dynamic qualitative phenomena and to represent stimulus-response interrelations of persons in a group. The symbol  $I^0$  represents a qualitative characteristic, and its time rate ( $T^{-1}$ ) indicates it is an event, action, behavior, or something-going-on-in-time. The variety of things-going-on in the picture are suggested, but without fuller sampling of them they are somewhat indefinitely delimited as presented in a still picture—therefore, the class script is underlined to denote indefiniteness.

Of the twenty human beings in the picture two are children responding to the sort of stimuli represented by the other eighteen. Sixteen of these eighteen are in one of seven interacting groups (pairs or triads). (The hatted man at middle right and the "fiend" at lower left are isolated in their situations as portrayed.) The persons in these groups are stimulating and responding to each other. The two children spectators are responding to a complex movie situation representatively portrayed in the picture. The children are stimulated by the individual actors and also by them in groups, i.e., by the situation of their interacting. All this is symbolized in the S-formula by the partial matrix of interacting persons and groups which in expanded form is of the following type (note carefully the empty cells):

	$P'$	$P''$	$P'''$	$P'''$		$ P'$	$ P''$	
$P'$	—							= girl spectator
$P''$		—						= boy spectator
$P'''$	$I_p^0T^{-1}$	$I_{p'}^0T^{-1}$	—	$I_{p''}^0T^{-1}$				= heroine at lower right
$P'''$	$I_{p'}^0T^{-1}$	$I_{p''}^0T^{-1}$	$I_{p'''}^0T^{-1}$	—				= villain at lower right
								the 16 other persons in turn
$ _p$	$I_{p'''}^0T^{-1}$	$I_{p''}^0T^{-1}$				—		= the heroine-villain pair at lower right
$ _{p''}$	$I_{p'''}^0T^{-1}$	$I_{p''}^0T^{-1}$					—	= another pair or triad
								the five other groups in turn

=  $p :: pI_{p/t}^0$   
in Brief-S  
(Eq. 107,  
Ch. X)

S. 53 (*Continued*)

The capital letter descript (each column) denotes a responder; the small letter descript (each row) denotes a stimulator. The primes on the attributes I<sup>0</sup> denote different particular qualitative responses. Thus, the two spectators are not stimulating nor responding to each other (in the S-situation as pictured) but are only responding to the actors and groups of them. Each actor is stimulating and responding to the other actor (or actors) in the same pair (or triad). Each actor and each group of actors is stimulating but not responding to each spectator, i.e., there is contact but not interaction between actors and spectators (as these terms are defined in Chapter VII). The expanded matrix compels exact analysis of the stimulators, responders, and the nature of the responses (I<sup>0</sup>T<sup>-1</sup>)<sub>1</sub>, and thus specifies the structure of a sociological group (which is defined as a plurel of interrelated parties).

## S. 54

## INCREASES IN RACIAL FRIENDLINESS

*Reported by a Case-Group of  
100 American Men*

*Toward*

Germans	21
Japanese	13
Swedish	11
Chinese	10
French	10
Italians	10
Spanish	10
Filipinos	8
Russians	8
English	7
Czecho-Slovaks	6
Indians (Amer.)	6
19 other races	54
TOTAL	174

*Reported by a Case-Group of  
100 American Women*

*Toward*

Japanese	25
Negroes	18
Jew-German	17
Chinese	16
Jew-Russian	13
Indians (Amer.)	11
Italians	10
Armenians	9
Mexicans	9
Canadians	8
Norwegians	8
Swedish	8
17 other races	44
TOTAL	196

*Ref.: Bogardus, Emory S., "Sex Differences in Racial Attitudes," Sociology and Social Research, Vol. XII, 1927-28, p. 285.*

*Descriptive formula:*  $S_{54} = P_p : \underline{IT}^{-1} : \underline{P}_q$

*Legend:*

$S_{54}$  = The situation  
records for

$P$  = 200 Americans  
in

$|_p$  = 2 sex plurels

*Quantic number* = 9;1;0;11

$\underline{I}$  = their increased friendliness  
(amount not stated)

$\underline{T}^{-1}$  = in an unspecified period  
towards each of

$\underline{P}_q$  = 19 racial plurels

*Comment:*

The graph illustrates an imperfect contacting matrix, as the increases of friendliness are all one-way changes of attitude of the Americans towards other races. Therefore, by Rule 36 (Appendix II) the quantic number is 11 in the populational sector, not 2, which would denote full two-way interaction.

## S. 55

## INDEX OF THE VOLUME OF ACQUAINTANCES OF 16 INDIVIDUALS

Name	I.Q.	Cottage	After 30 Days	After 60 Days	After 90 Days	After 120 Days	After 150 Days	After 180 Days
JN	100	C14	13	18	33	39	41	42
GU	121	C4	63	65	42	26	29	28
RD	62	C4	7	8	12	9	9	8
DB	85	C6	30	43	42	46	73	72
ML	80	C4	24	27	30	33	27	28
MK	86	C2	10	12	25	38	29	30
SO	112	C4	30	44	37	50	62	74
KN	87	C9	21	32	33	52	101	131
IL	116	C4	42	61	50	28	46	43
DN	85	C6	22	42	29	32	34	31
HY	102	C8	15	12	24	31	51	46
HR	65	C14	9	9	10	11	14	13
RZ	88	C14	33	14	22	25	25	26
HF	91	C8	30	44	79	84	75	82
FA	77	C8	14	16	15	32	32	33
NI	82	C9	13	25	41	42	47	49

Ref.: Moreno, J. L., *Who Shall Survive?* Nervous and Mental Disease Publishing Co., 1934, p. 137.

Descriptive formula:  $S_{55} = {}^pP_{,p} : I : {}^tT^{-1} : P$

Quantic number = 9;1;0;11

Legend:

$S_{55}$  = The situation

$I = IQ's$

records for each of

and, on each of

${}^pP$  = 16 named persons

${}^tT^{-1}$  = 6 dates, 30 days apart,

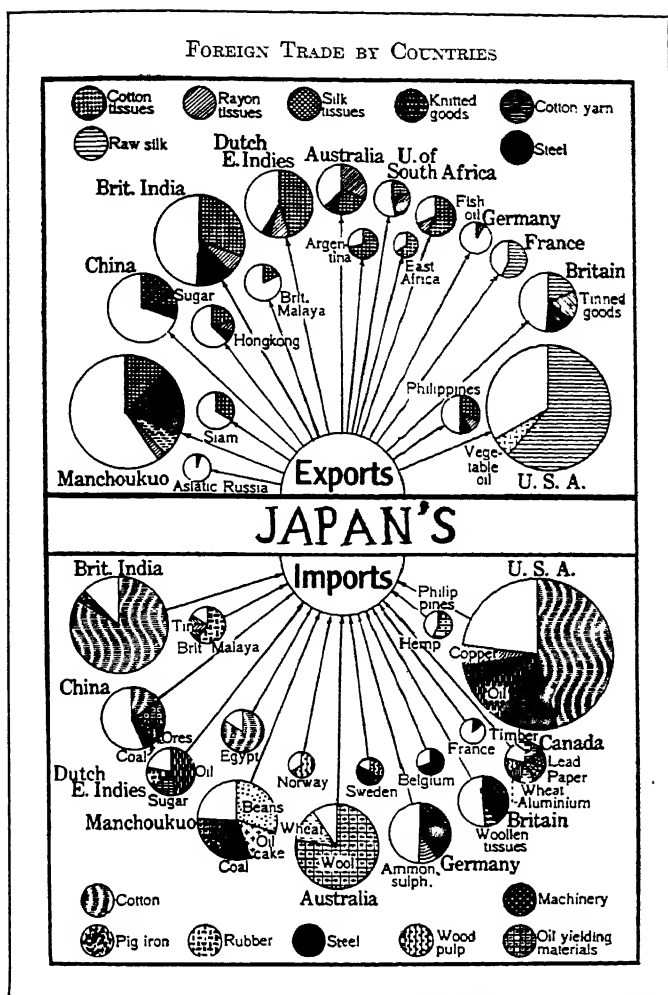
$|_{,p}$  = their Cottage plurel

$P$  = the number of their acquaintances

and

Comment:

Contacting, i.e., change of one-way relations, is illustrated. If the persons claimed as acquaintances had recorded their acquaintances, the contacting would be interacting, i.e., change of the two-way relations between persons. Again note that by Rule 36 (Appendix II) the quantic number of contact is 11, instead of the 2 denoting interaction.

*S.* 56

Ref.: Yano, T. and Shirasaki, K., *Nippon, A Chartered Survey of Japan*, Tokyo, 1936, p. 47.

*Descriptive formula:*  $S_{56} = P, :: P_p : t'(\sigma IT^{-1})_i$

Quantic number = 9;1;0;2

*Legend:*

S<sub>56</sub> = The situation  
records

P, = Japan

$::$  = interacting with

$P_p = 23$  nations

in

%I = percentage trade

 $\tau^{-1} = \text{per year (in 1935)}$ 

in each of

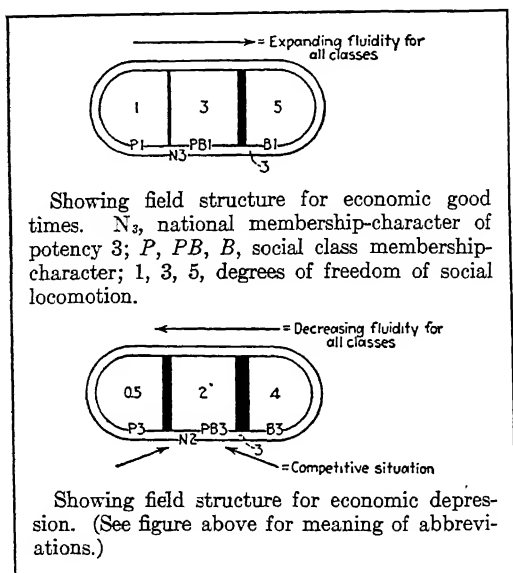
 $I_i = 13$  commodities

## S. 56 (Continued)

## Comment:

This graph in matrix terms is a single row and column of the economic interaction matrix with percentages as the cell entries. It is developed into a third-degree matrix by the expansion for different commodities. Note that imports and exports are the two triangles of cells on either side of the main diagonal of the matrix.

## S. 57



Ref.: Brown, J. F., *Psychology and the Social Order*, McGraw-Hill, 1936, p. 136.

Descriptive formula:  $S_{57} = \underline{P}_q : \underline{a} : \underline{t}T^{-1} : ( : \underline{P}_q : \underline{I}_i, \underline{I}_i )$

Quantic number = 9;1;0;2

## Legend:

$S_{57}$  = The situation  
records for each of

$\underline{P}_q$  = a national

and

$\underline{a}$  = 3 economic subplurels

in each of

$\underline{t}T^{-1}$  = 2 periods (good times and depressions)

$\underline{P}_q$  = an inter-plurel

$\underline{I}_i$  = index of "permeability" of  
barriers (amount not stated)

and

$\underline{I}_i$  = indices of 2 characteristics of  
"fluidity" and "potency"

$\underline{I}_i$  = the indices being in ordinal  
units (i.e., relative ranks,  
"non-metricized")

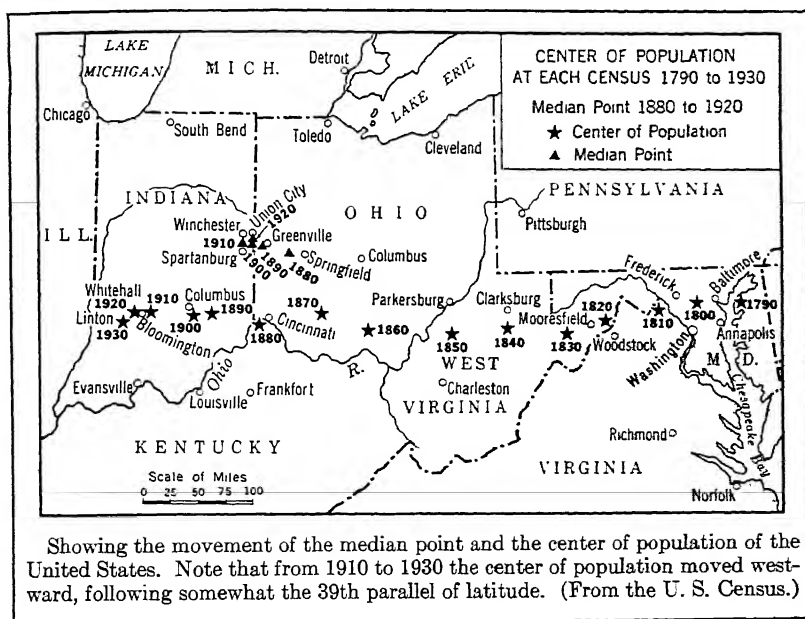
## S. 57 (Continued)

## Comment:

The diagrams represent change of static interrelations more than interacting proper. They also illustrate the combination in one situation of simple characteristics of plurels ( $I_1$ ) with interrelating characteristics ( $I_1$ ) between plurels.

The diagram is a typical illustration of topology, a field of qualitative mathematics or non-quantitative geometry, applied to societal data. The notation of S-theory can express topological data as well as metricized data.

## S. 58



Ref.: Huntington & Carlson, *Geographic Basis of Society*, Prentice-Hall, 1934, p. 22.

Descriptive formula:  $S_{ss} = {}^tT^{-1} : \underline{I}^2 : \underline{P}$

Quantic number = 9;0;2;1

Legend:

$S_{ss}$  = The situation  
records

${}^tT^{-1}$  = 15 decennial dates  
and corresponding  
to each

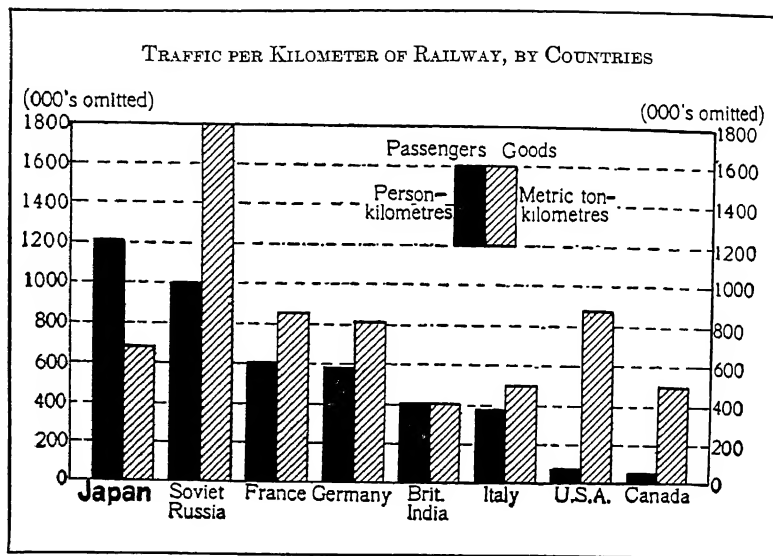
$\underline{I}^2$  = 2 mapped points { center  
and  
median

$\underline{P}$  = of the U.S. population

## Comment:

The situation illustrates the process of relocating, the moving of a point in an area. By measuring the distance which the center of population moves in a decade, "lineating" is determined, and the sigma of these distances determines "linear revarying," which is a measure of the constancy of the westward trend.

## S. 59



Ref.: Yano, T. and Shirasaki, K., *Nippon, A Chartered Survey of Japan*, Tokyo, 1936, p. 347.

Descriptive formula:  $S_{ss} = (P, I)_p LT^{-1}$

Quantic number = 9;1;1;1

Legend:

$S_{ss}$  = The situation  
records for each of

$T^{-1}$  = some unspecified period  
of

$|_p$  = 8 nations

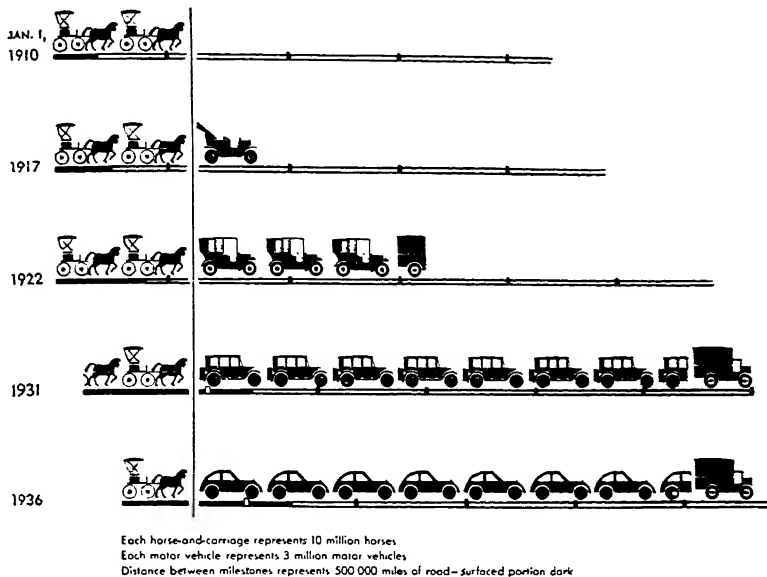
P = passengers

L = the kilometrage  
for

, = compared with  
I = tons of goods

S. 60

## MOTOR VEHICLES AND PUBLIC ROADS



Ref.: Pictorial Statistics, Inc., New York.

Descriptive formula:  $S_{60} = {}^tT^{-1} : (L_1, I_1)$ 

Legend:

Quantic number = 9;1;1;0

$S_{60}$  = The situation  
records for each of

 ${}^tT^{-1}$  = 5 dates, 1910-36,

L = the road length

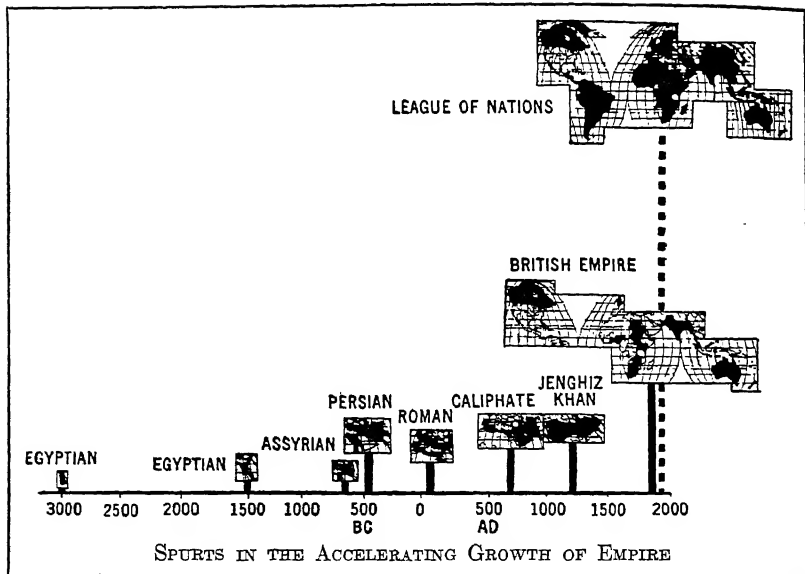
in

 $|_1$  = paved and unpaved classes

and  
I = the number of vehicles  
in

$|_1$  = 2 classes {  
carriages  
and  
motors

## S. 61



Ref.: Hart, Hornell, *The Technique of Social Progress*, Holt, 1931, p. 449.

Descriptive formula:  $S_{61} = {}^tT^{-1} : L_t^2$

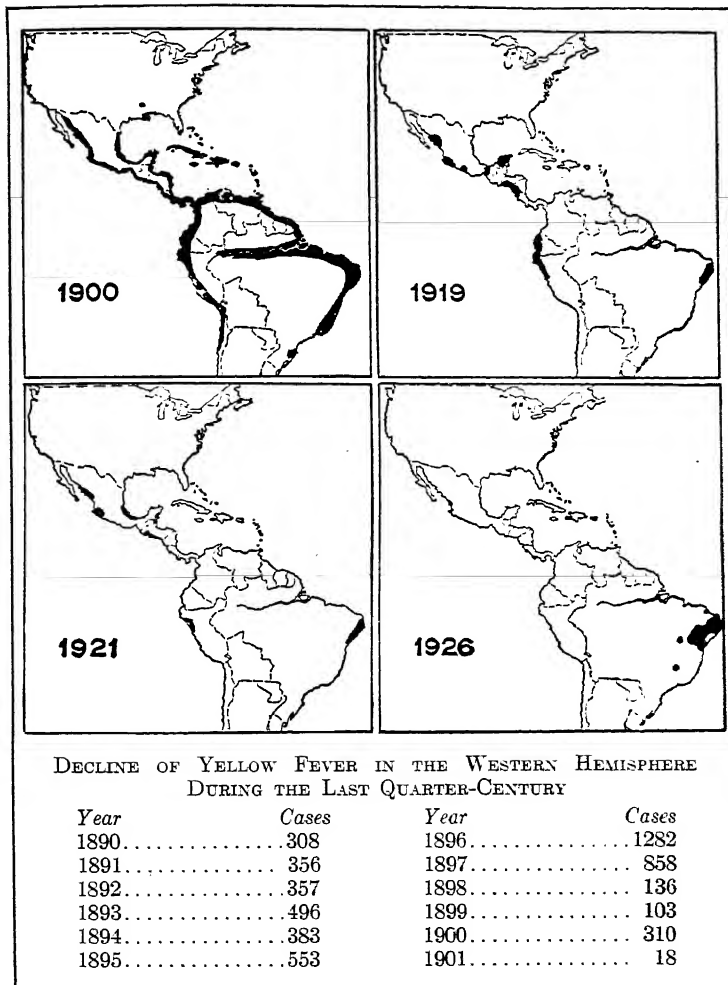
Quantic number = 9;0;2;0

Legend:

$S_{61}$  = The situation  
records for each of

${}^tT^{-1}$  = 9 dates  
 $L_t^2$  = an area of empire

S. 62



Ref.: Kelsey, Carl, *The Physical Basis of Society*, Appleton, 1928, p. 219.

## S. 62 (Continued)

Descriptive formula:  $S_{62A} = \text{tT}^{-1} : L_1^2$ 

Quantic number = 9;0;2;0

Legend:

 $S_{62A}$  = The situation

records for each of

 $\text{tT}^{-1} = 4$  years $L_1^2 = 2$  zones in the Americas  
{ yellow fever  
} and noneDescriptive formula:  $S_{62B} = \text{tT}^{-1} : (PT^{-1})$ 

Quantic number = 8;0;0;1

Legend:

 $S_{62B}$  = The situation

records for each of

 $\text{tT}^{-1} = 12$  years

'| = beginning in 1890

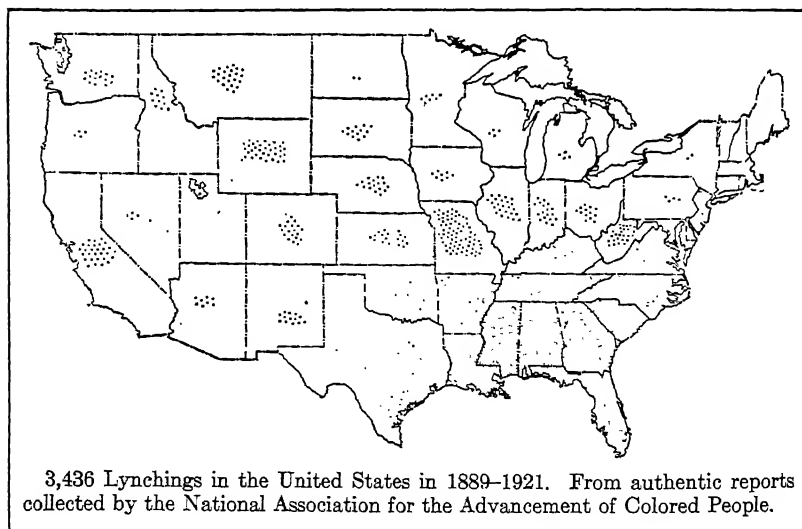
P = the cases of yellow fever

 $T^{-1}$  = per year

Comment:

Two different quantics are involved in the map and in the table, as one has  $L^2$  and the other  $L^0$ , and the two L arrays are not contiguous. Therefore, there are technically two S-situations presented in this excerpt from page 219 of Kelsey's book.

## S. 63



Ref.: Survey, Vol. XLVII, No. 24, March 11, 1922, p. 916.

Descriptive formula:  $S_{63} = \underline{L}_1^2 : \text{t}(PT^{-1})$ 

Quantic number = 9;0;2;1

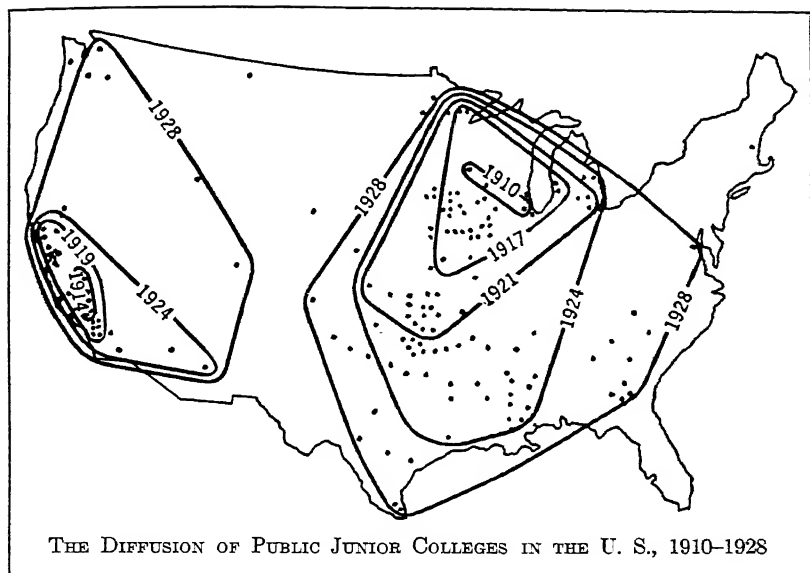
Legend:

 $S_{63}$  = The situation

records for each of

 $\underline{L}_1^2 = 48$  statesP = the number of lynchings  
in the $\text{tT}^{-1} = 1889-1921$  period

S. 64



Ref.: Pemberton, H. Earl, "The Special Order of Culture Diffusion," *Sociology and Social Research*, Jan.-Feb., 1936, Vol. XXII, No. 3, p. 250.

Descriptive formula:  $S_{64} = (^{\circ}T^{-1} : ^1L^2 : ^1:P)_1$       Quantic number = 9;0;2;1

Legend:

$S_{64}$  = The situation

$L^2$  = a zone

records for each of

and

$|_1$  = 2 diffusion areas in the United States

$|$  = point locations

$^1|$  = each with a corresponding

for each of

$P$  = Junior College plural

$^{\circ}T^{-1}$  = five dates

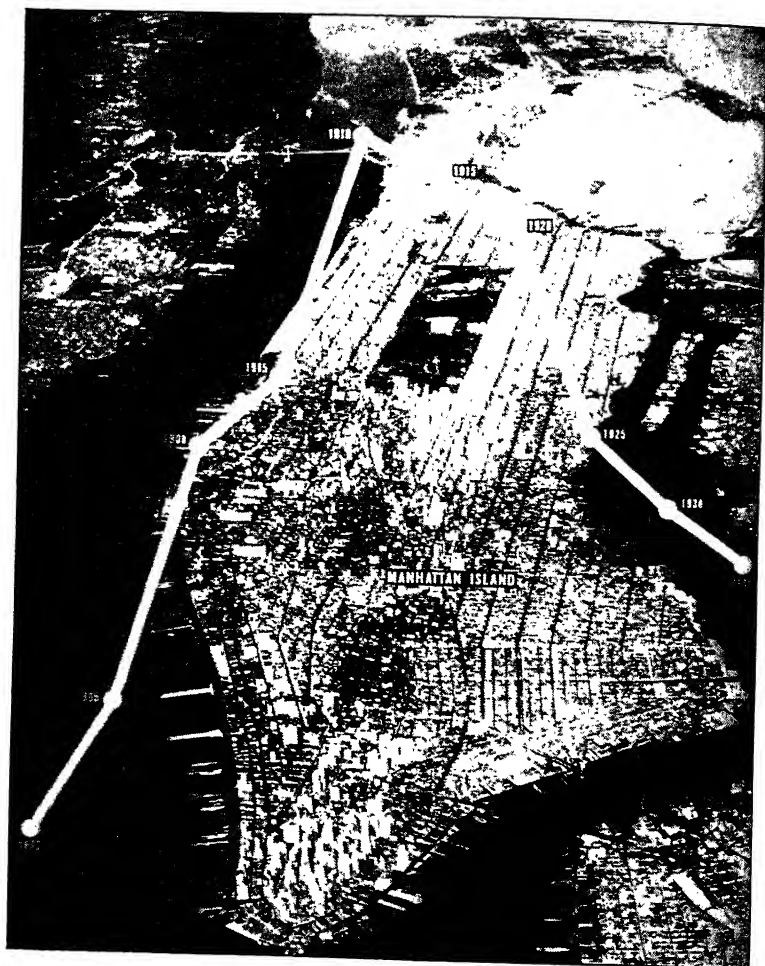
Comment:

The situation maps the temporal and spatial diffusion of a culture complex.

Comment on notation:

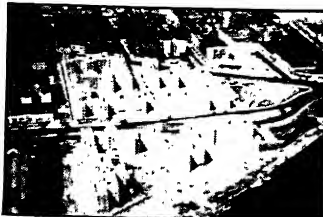
Since the areal index has both areal and point scripts ( $|$ ), it would be ambiguous as to whether there was a plural ( $P$ ) corresponding to each of the five areas ( $^{\circ}T^{-1} : L^2$ ) or to each point ( $^1L^2$ ), unless the point script were repeated as a cross script on the  $P$ , thus ( $^1:P$ ).

S. 65



When Manhattan's population started rising, as shown by the graph above, the roof seemed the limit. The roofs continued to rise till 1929. But population itself began to tumble around 1910 and now has slipped back 30 years. It is still going down, partly because people are spilling out into the suburbs and partly because the human supply has been shut off by immigration restrictions and birth control. The later growth of skyscrapers intensified congestion, increased traffic snarls, added to the cost of doing business, raised rents, helped create blighted districts.

Large-scale housing demands low land values. Hence it has begun on the blighted edges of the city where values are low. The PWA Harlem Houses (*right*) show what can be done to rebuild on an open pattern, proving that you don't have to become a farmer or a suburbanite to have sunlight, air, recreation areas. If the big city is to hold its own it must provide the equivalent of Harlem Houses for all.



S. 65 (*Continued*)

*Descriptive formula:*  $S_{65} = \underline{L}^2 : ({}^tT^{-1} : P) : I_v^0.$       *Quantic number* = 9:0:2:1  
*Legend:*

$S_{65}$  = The situation  
           records

${}^t|$  = beginning in 1890

$P$  = the population

$\underline{L}^2$  = an area (Manhattan)  
           and corresponding to it,  
           for each of

and applied to this area

$I_v^0$  = an evaluative attribute  
           ("Bad")

${}_t|$  = 9 periods

$T^{-1}$  = of five years each

*Comment:*

The picture is one of a series on city layouts, illustrating types labeled "Good" and "Bad" in the author's judgment. Note that the analyst writing the S-formula does not introduce his evaluation; he merely records in sixteen standardized S-symbols what is recorded in all sorts of other symbols (a word, a curve with numbers, and an air-photo-map, here) in the situation as presented for analysis.

## BETWEEN ROME AND CHINA

The second and first centuries B.C. mark a new phase in the history of mankind. Mesopotamia and the eastern Mediterranean are no longer the center of interest. Both Mesopotamia and Egypt were still fertile, populous and fairly prosperous, but they were no longer the dominant regions of the world. Power had drifted to the west and to the east. Two great empires now dominated the world, this new Roman Empire and the renascent Empire of China. Rome extended its power to the Euphrates, but it was never able to get beyond that boundary. It was too remote. Beyond the Euphrates the former Persian and Indian dominions of the Seleucids fell under a number of new masters. China, now under the Han dynasty, which has replaced the Ts'in dynasty at the death of Shi-Hwangti, had extended its power across Tibet and over the high mountain passes of the Pamirs into western Turkestan. But there too it reached its extremes. Beyond was too far.

China at this time was the greatest, best organized and most civilized political system in the world. It was superior in area and population to the Roman Empire at its zenith. It was possible then for these two vast systems to flourish in the same world at the same time in almost complete ignorance of each other. The means of communication both by sea and land were not yet sufficiently developed and organized for them to come to a direct clash.

Yet they reacted upon each other in a very remarkable way, and their influence upon the fate of the regions that lay between them, upon central Asia and India, was profound. A certain amount of trade trickled through, by camel caravans across Persia, for example, and by coasting ships by way of India and the Red Sea. In 66 B.C. Roman troops under Pompey followed in the footsteps of Alexander the Great, and marched up the eastern shores of the Caspian Sea.

Ref.: Wells, H. G., *A Short History of the World*, Watts and Co., 5 and 6 Johnson's Court, Fleet Street, E.C. 4, London, 1929, p. 126.

*Descriptive formula:*  $S_{66} = {}_tT^{-1} : P_p : \underline{L}^2 : {}^tI_1^{9,1}$

*Quantic number* = 9;1;2;1

*Legend:*

$S_{66}$  = The situation  
records in  
 ${}_tT^{-1}$  = the first 2 centuries B.C.  
for each of  
 $P_p$  = 2 empires, Roman and Chinese  
 $\underline{L}^2$  = their territory  
and

$I_1$  = 15 other characteristics

{ "interest"  
"fertile"  
"power"  
"civilized"  
"trade," etc.

$|^0$  = both qualitative

and

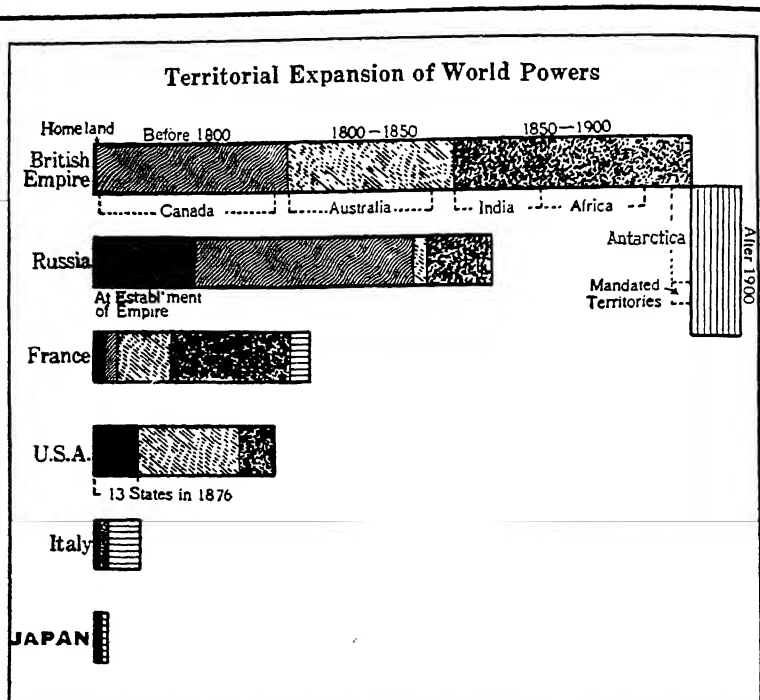
$|^{+1}$  = ordinally quantitative  
("best," "most," etc.)

*Comment:*

The prose paragraph is presented as an illustration which is on the borderline of being a quantitatively recorded situation such as S-theory applies to. The

data are vague, since the periods are not sharply defined, nor the geographic boundaries and area mapped; and the characteristics denoted by such words as "greatest," "best organized," "most civilized," lack objective indicators. Although predominantly qualitative, incipient quantifying of these adjectives is denoted by the superlative degree in "greatest," "best," "most," etc. The superlative degree converts an attribute into a point on an ordinal indicant,  $I^{+1}$ . The analysis into S-notation brings into clear relief the degree of indefiniteness and inexactness in the usual narrative style of historical writing. Literary style is, of course, superior if the historian's purpose is to interest the lay reader. But, if it is to record events as facts for the inductions of the social scientists, definiteness and completeness, promoted by expression in the matrix formulae of S-theory, become essential.

## S. 67

**Territorial Expansion of World Powers**

Square kilometres (000's omitted)

	Home land	Territories held in the indicated years				
		Before 1800	1800-1850	1850-1900	After 1900	Present Totals
Japan .....	382	—	—	36	1) 263	681
Brit. Empire .....	243	10 248	8 848	12 410	7 926	39 678
France .....	551	721	2 838	6 422	1 052	11 584
Italy .....	310	—	—	365	1 755	2 430
Russia .....	2) 5 392	11 775	638	3 538	—	21 345
U.S.A. ....	3) 2 310	—	5 463	1 919	2	9 694

1) Including Kwantung Leased Territory & the Mandated South Seas Islands.

2) As at the end of 15th century when an independent Empire.

3) 13 States as at the Declaration of Independence.

## S. 67 (Continued)

Descriptive formula:  $S_{67} = \underline{P}_p : {}^aT^{-1} : L_1^2$ 

Quantic number = 9;0;2;1

Legend:

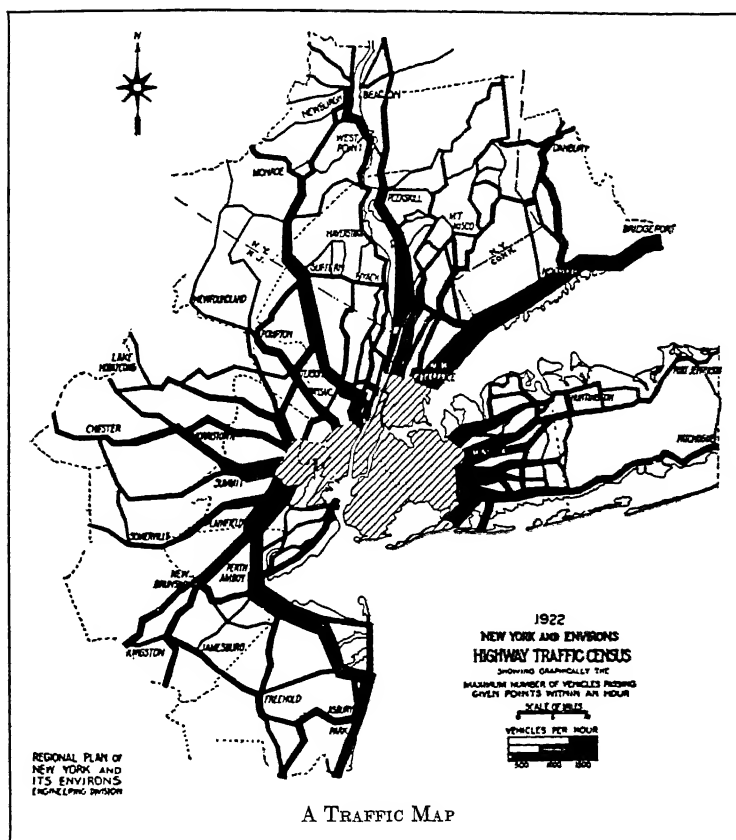
$S_{67}$  = The situation  
records for each of

 ${}^aT^{-1}$  = 6 periods, $^a|$  = with varying initial dates, $\underline{P}_p$  = 6 nations $L^2$  = increases of area

in each of

 $|_1$  = of homeland and territories

## S. 68



Ref.: Arkin, H., and Colton, R. R., *Graphs: How to Make and Use Them*, Harpers, 1936, p. 149. From Lewis, H. M., and Goodrich, E. P., *Regional Survey of New York and Its Environs*.

S. 68 (*Continued*)

*Descriptive formula:*  $S_{68} = \underline{L}_1^2 : \underline{L}_1^{\tau 1} : \tau (IT^{-1})$   
*Legend:*

$S_{68}$  = The situation  
 records

$\underline{L}_1^2$  = in the counties around New  
 York City

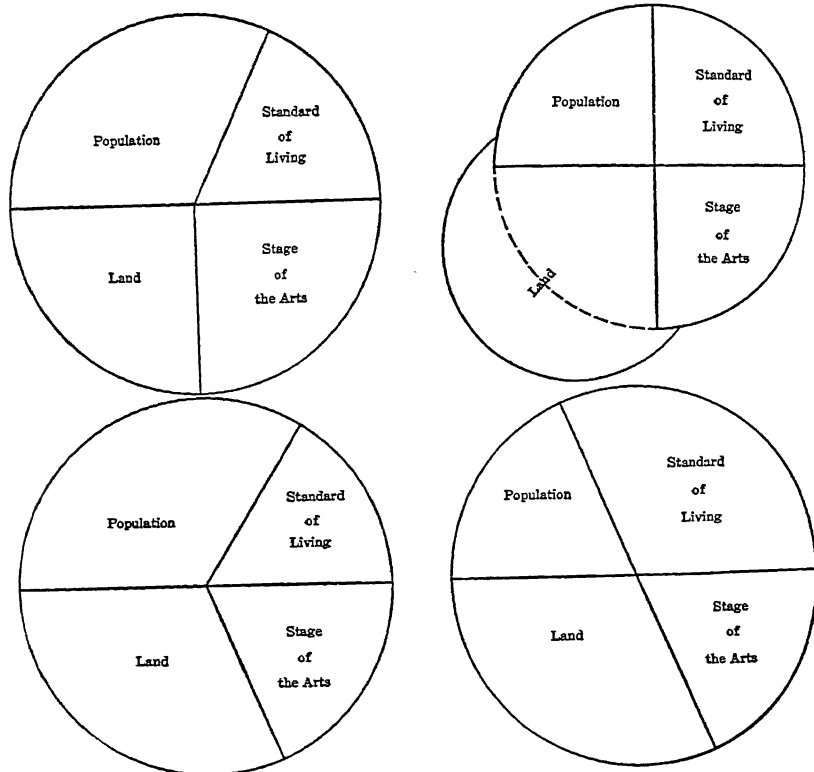
$\underline{L}_1^{\tau 1}$  = the trunk roads

with  
 $I$  = the number of cars passing  
 $\tau (T^{-1})$  = per hour in 1922

*Comment on notation:*

Note the double quantic digits of 21 in the space sector, which by Rule #36 (Appendix II) denote the aggregate of areas and lines in the situation as recorded.

## S. 69



S. 69 (*Continued*)*Descriptive formula:* S<sub>69</sub> = <sup>v</sup>T<sup>-1</sup> : P : L<sup>2</sup> : I<sub>i</sub>*Quantic number* = 9;1;2:1*Legend:*

S <sub>69</sub> = The situation	and
records for each of	<u>I</u> = indicants (with indefinite
<sup>v</sup> T <sup>-1</sup> = 4 dates,	units)
a corresponding	of
P = population	{ standards of
and	
L <sup>2</sup> = area,	<sub>i</sub> = 2 kinds { living,
	technology

*Comment:*

Strictly, the situation compares four societies and represents a change only if, as implied in the text, the four circles represent one society on four dates. The latter interpretation is taken for writing the quantic.

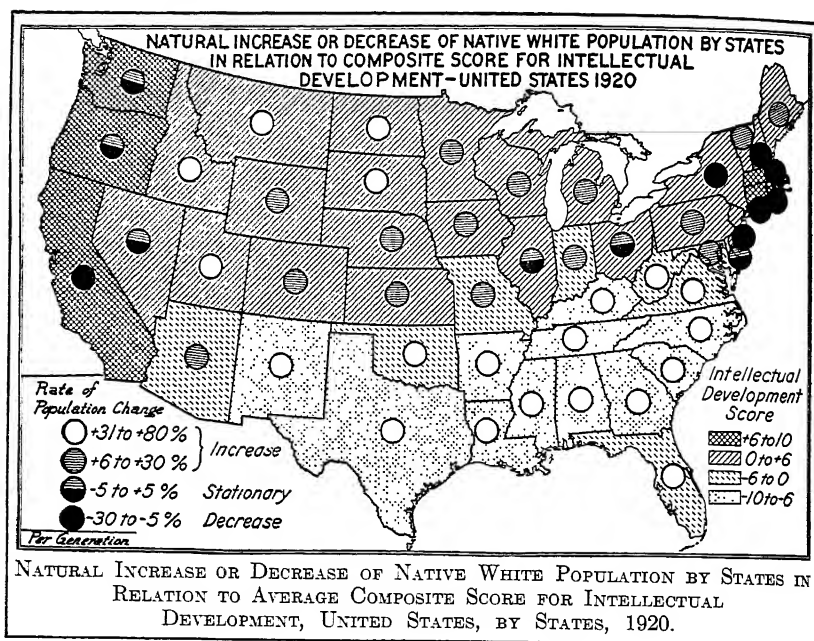
This analysis asserts a correlation between the standard of living, I<sub>,,</sub>, and the three components of population, P, technology ("stage of the arts"), I<sub>,,,</sub>, and land. Land in S-theory would be subdivided into area, L<sup>2</sup>, and natural resources, I<sub>,,,</sub> (i.e., a composite index of fertility, minerals, rainfall, etc.). The hypothesis in Fairchild's diagrams above are then expressible in S-notation, which facilitates quantitative verification, as:

$$I, \bullet (P_{,,}; I_{,,}; I_{,,,}; L^2) = r > 0 \quad (\text{Eq. 108, Ch. X})$$

This implies that I, correlates with each of the four indices in parenthesis, and therefore, must correlate more highly with some optimal combination of them. This optimal combination may be calculated, as a first approximation, by multiple correlation technics as a weighted sum. As ratios, products, etc. may prove to fit the data more closely, the semicolon denoting any particular but unspecified operator combining them is written here.

This economic analysis may be compared to the classical analysis whereby land, labor, and capital are considered to be the chief factors in production. (See S. 25, Ch. II for an equation relating these.) Denoting the number of laborers by P<sub>,,,</sub>, capital by I<sub>,,,</sub>, these indices can be included in Eq. 108, Ch. X. Overlapping of the indices is measured by their intercorrelations. Component analysis of the intercorrelational matrix, as described in Chapter VI, should reveal a minimal set of non-overlapping components which most completely determine the standard of living.

## S. 70



Ref.: Lorimer and Osborn, *Dynamics of Population*, Macmillan, 1934, p. 183.

Descriptive formula:  $S_{70} = 'T^0 : L_1^2 : (I, \%PT^{-1})$       *Quantic number* = 9;1;2;1

*Legend:*

$S_{70}$  = The situation  
records

I = an intellectual score  
and

$'T^0$  = in 1920  
in each of

$\%P$  = the percentage population  
change

$L_1^2$  = the 48 states

$T^{-1}$  = per generation

*Comment on notation:*

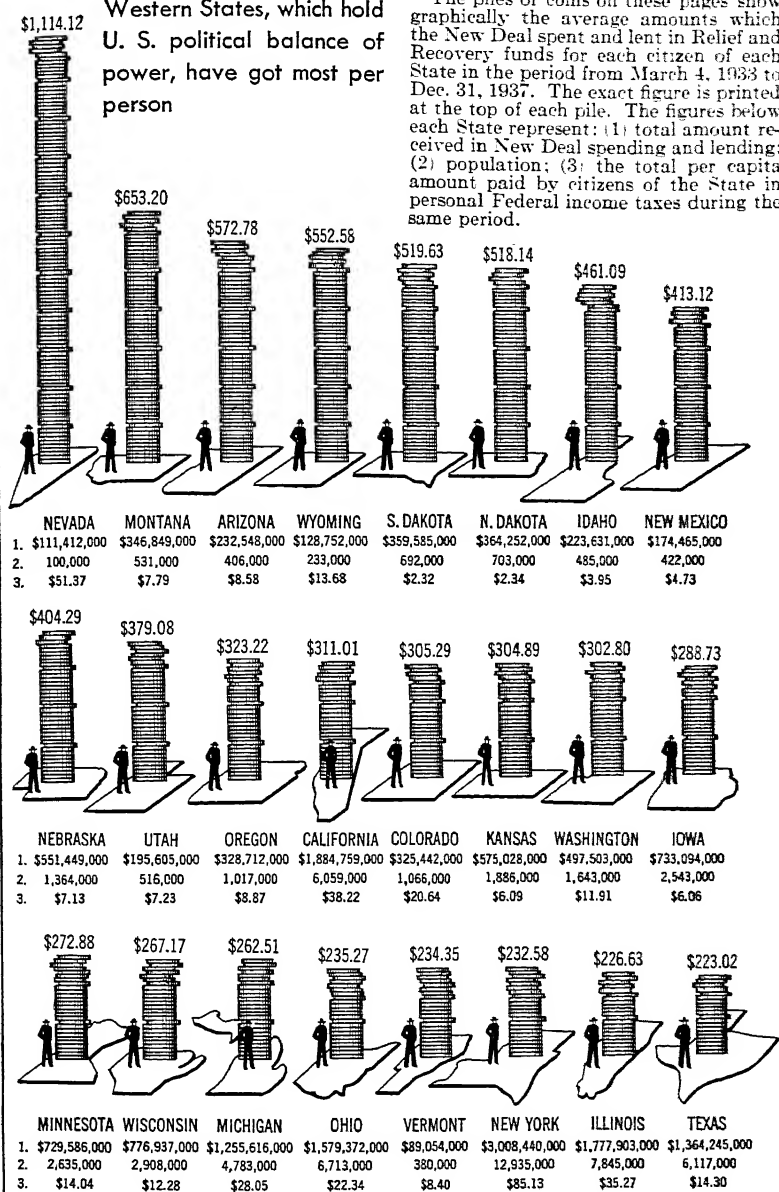
The implied correlation between intellect and population change has not been operationally calculated and hence does not require the quantic of a correlation,  $I^2$ .

## S. 71

## HAS THE NEW DEAL SPENT BILLIONS TO KEEP ITSELF IN POWER?

Western States, which hold U. S. political balance of power, have got most per person

The piles of coins on these pages show graphically the average amounts which the New Deal spent and lent in Relief and Recovery funds for each citizen of each State in the period from March 4, 1933 to Dec. 31, 1937. The exact figure is printed at the top of each pile. The figures below each State represent: (1) total amount received in New Deal spending and lending; (2) population; (3) the total per capita amount paid by citizens of the State in personal Federal income taxes during the same period.



S. 71 (*Continued*)

*Descriptive formula:*  $S_{71} = \underline{L}_1^2 : {}_t({}_sIP^{-1}T^{-1})_i$       *Quantic number* = 9;1;2;9

*Legend:*

$S_{71}$  = The situation  
records for each of

$\underline{L}_1^2$  = 24 of the United States  
for

${}_tT^{-1}$  = the period March, 1933 to  
December, 1937

$( )_i$  = two financial indices in  
dollars

namely

$({}_sIP^{-1})_i$  = (1) income taxation per  
capita

and

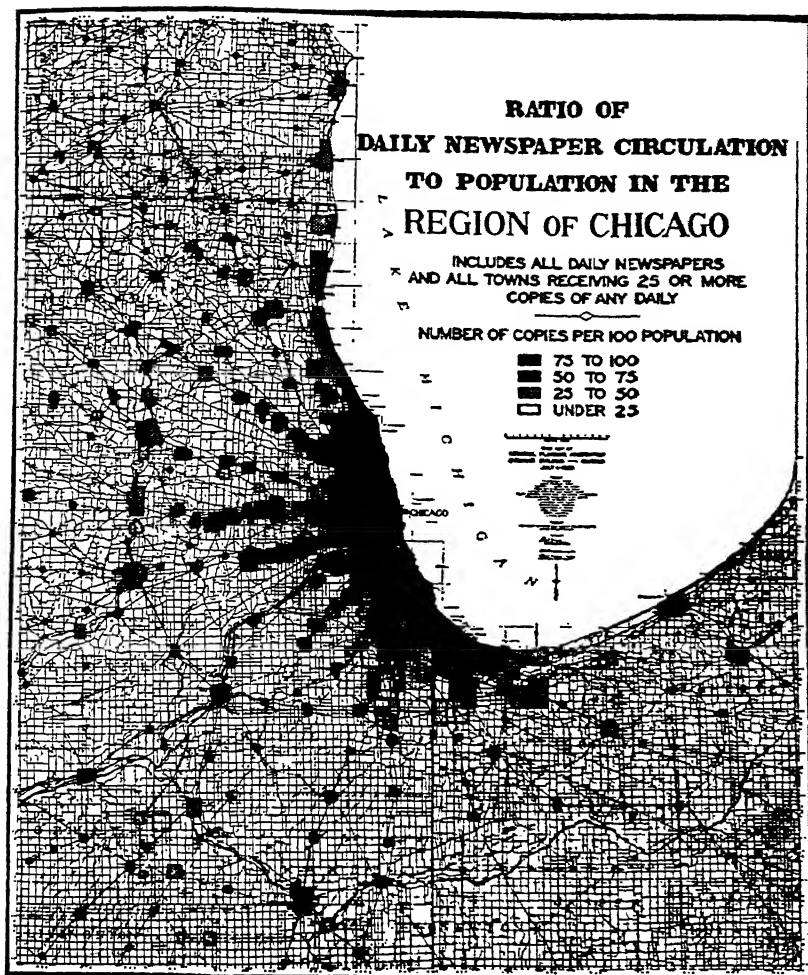
$I_{,,}$  = (2) the Government ex-  
penditure,

$P_{,,}$  = the population,

and

$(IP^{-1})_{,,}$  = the per capita expenditure

S. 72



Ref.: Parks, Robert E., "Urbanization as Measured by Newspapers Circulated," *American Journal Sociology*, Vol. XXXV, No. 1, 1929, p. 66.

Descriptive formula:  $S_{72} = \underline{L}_1^2 : (IP^{-1}T^{-1})$

Quantic number = 9;1;2;9

Legend:

$S_{72}$  = The situation  
records for every one of

$\underline{L}_1^2$  = Chicago's districts

I = the number of copies of  
newspapers

$P^{-1}T^{-1}$  = per person per day

# DISTRIBUTION OF ANSWERS TO QUESTIONING (No.44) ABOUT AMOUNT OF WATER Maximum score = 30

The "normal" rural Arab sample (N = 100) (families identified by number)  
compared with the urban sample (N = 50) (each family indicated by two commas)

Liters of Water per Person per Day																																		
Sample (N = 50) (each family measured by two commas)																																		
0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200	210	220	230	240	250	260	270	280	290	300	Rural	Urban		
31	01 04 13 48 50	52 61 67 72 77	79 83 84 90 99 00	30 33 43 44 45	46 47 51 53 56	57 58 59 60 63 68 69	70 71 73 74 76 78 80 81	82 85 87 88 89	01 03 04 95 97																									
0	01 04 13 48 50	52 61 67 72 77	79 83 84 90 99 00	30 33 43 44 45	46 47 51 53 56	57 58 59 60 63 68 69	70 71 73 74 76 78 80 81	82 85 87 88 89	01 03 04 95 97																									
10-19	02 03 05 07 08 09 10 11 12 14 15 16 18 20 29 30 33 43 44 45 46 47 51 53 56																																	
20-29	06 17 22 24 34 35 37 39 41 62 04 65 75 86 92 96 98																																	
30-39	10 23 25 26 28 32 36 38 40 42 46 66																																	
40-49	27 54																																	
50-59	55																																	
60-69																																		
70-79																																		
80-89																																		
90-99	21																																	
100-109																																		
110-119																																		
120-129																																		
130-139																																		
140-149																																		
150-159																																		
160-169																																		
170-179																																		
180-189																																		
190-199																																		
200-209																																		
210-219																																		
220-229																																		
230-239																																		
240-249																																		
250-259																																		
260-269																																		
270-279																																		
280-289																																		
290-299																																		
300-309																																		

*Comment.*  
Perhaps the most outstanding rural-urban difference is in the amount of water used in the home. The city house which uses the least water gets as much as the village house which uses the most. There is almost no overlap in the graph of the two distributions. This is the most important cause of differences in cleanliness, hygiene, and consequent health.  
The village water is brought in jars on the woman's head from a well or spring. Frequently two hours a day are required for the woman to go, draw the water, and return. Naturally water is used sparingly, the average amount being about 19 liters a day per person (about one kerosene tin). Families using as little as 3 liters a day per person were found!  
In the city there is running piped water in almost every house. Water meters are not provided for less than 250 liters a day per house. The result is that rural consumption per person is only about one sixth the urban consumption.  
On the dry plains and hills of Syria the water supply is basic to both health and wealth.

## Comment.

Perhaps the most outstanding rural-urban difference is in the amount of water used in the home. The city house which uses the least water gets as much as the village house which uses the most. There is almost no overlap in the graph of the two distributions. This is the most important cause of differences in cleanliness, hygiene, and consequent health.

The village water is brought in jars on the woman's head from a well or spring. Frequently two hours a day are required for the woman to go, draw the water, and return. Naturally water is used sparingly, the average amount being about 19 liters a day per person (about one kerosene tin). Families using as little as 3 liters a day per person were found!

In the city there is running piped water in almost every house. Water meters are not provided for less than 250 liters a day per house. The result is that rural consumption per person is only about one sixth the urban consumption.

On the dry plains and hills of Syria the water supply is basic to both health and wealth.

Ref.: Dodd, Stuart C., *A Controlled Experiment on Rural Hygiene in Syria*, American Press, Beirut, 1934, Table 42.

## S. 73 (Continued)

*Descriptive formula:*  $S_{73} = (M, \sigma, L^3 P^{-1} T^{-1} : P)_q$       *Quantic number* = 9;0;3;1  
*Legend:*

$S_{73}$  = The situation

records for each of

$q = 2$  density plurals

$L^3$  = the liters of water used

$P^{-1}$  = per person

$T^{-1}$  = per day

with corresponding

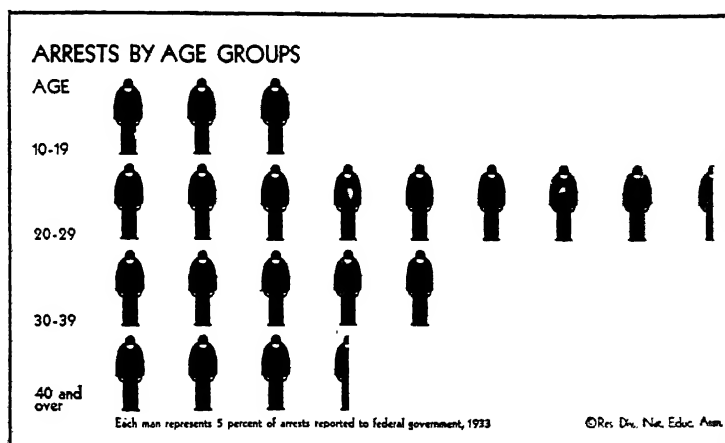
$P$  = listed frequencies of families

$M, \sigma$  = the mean and sigma are also stated

*Comment:*

The process of volumating is measured by the liters per day, and the process of revolving is measured by the sigma (S.D.). Revolving here is a spatial subform of reordinating, the dispersion of the volume of water used daily among the families.

## S. 74



*Ref.:* Research Bulletin of the National Education Association, Vol. XII, No. 5, Nov., 1934, p. 274.

*Descriptive formula:*  $S_{74} = tT^{+1} : (\%P, T^{-1})$       *Quantic number* = 19;0;0;1  
*Legend:*

$S_{74}$  = The situation

records

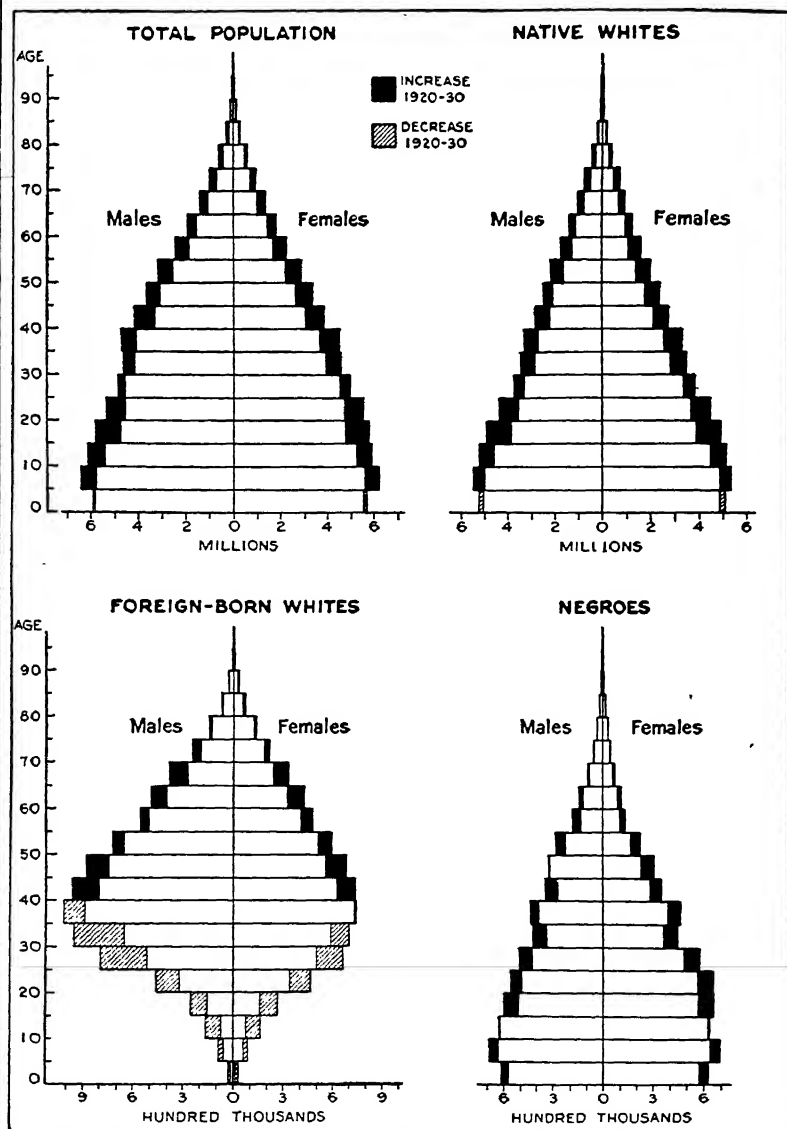
$tT^{+1}$  = 4 age intervals

with a corresponding

$\%P$  = % of arrests

$T^{-1}$  = for the year 1933

S. 75



Distribution by five-year age periods of the total population, native whites, foreign born whites and Negroes, 1920-1930.

S. 75 (*Continued*)

*Descriptive formula:*  $S_{75} = ({}^tT^{-1} : {}_tT^{+1} : P_p)_{q,\Sigma q}$       *Quantic number* = 91;0;0;1  
*Legend:*

$S_{75}$  = The situation

records for each of

${}^tT^{-1}$  = 2 dates (1920 and 1930)

${}_t|$  = for 20 age intervals

$T^{+1}$  = of 5 years each

$P$  = the U.S. population

$|_p$  = by sexes

for each of

$|_q$  = 3 racial plurels

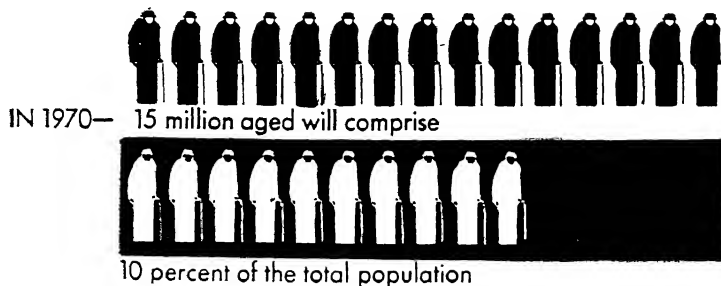
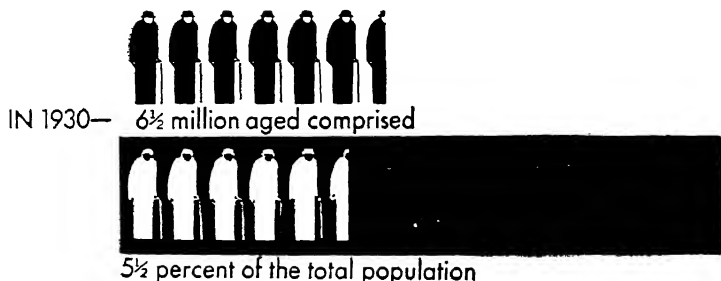
$|\Sigma q$  = and their sum

*Comment on notation:*

The parenthesis shows that the whole changing population pyramid is repeated for each of the  $|_q$  plurels. The change ( $T^{-1}$ ) precedes the duration ( $T^{+1}$ ) in the formula as the change dominates, i.e., the situation is one of changing ages, not one of aging changes. The processes of durational adpopulating, durating (rising average age), and of skewing in the foreign-born plurel can be measured from the data presented here.

S. 76

## PERSONS 65 AND OVER (IN THE UNITED STATES)



Descriptive formula:  $S_{76} = {}^tT^{-1} : {}^tT^{+1} : {}^cP$

Quantic number = 91 ; 0 ; 0 ; 1

Legend:

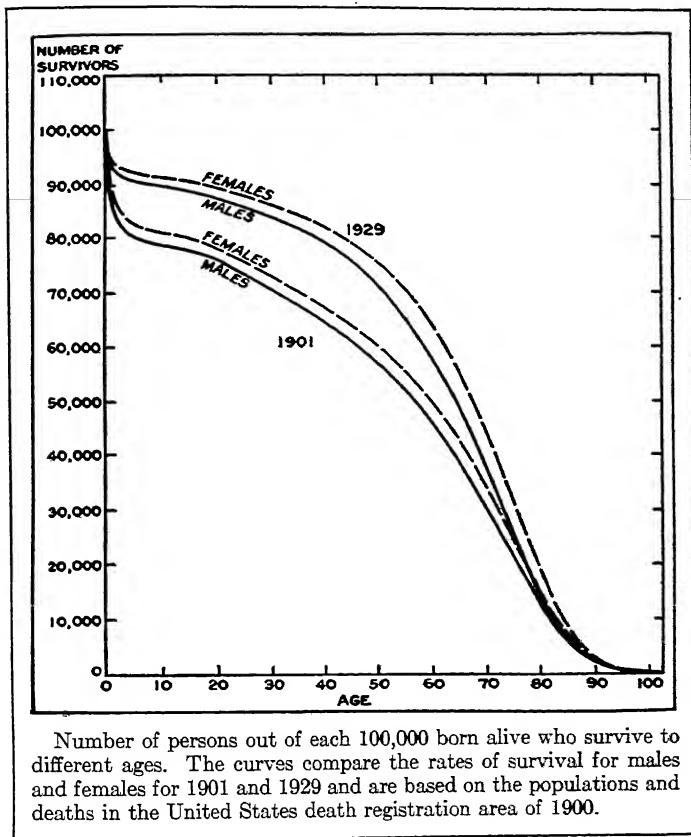
$S_{76}$  = The situation  
records for each of

${}^tT^{-1}$  = 3 dates (1890, 1930, 1970)  
for

${}^tT^{+1}$  = the age interval of 65 and  
over

${}^cP$  = the percentage of such per-  
sons in the population

S. 77



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, Vol. I, 1933, p. 606.

Descriptive formula:  $S_{77} = {}^tT^{-1} : {}^tT^{+1} : P_p$

Quantic number = 91 ; 0 ; 0 ; 1

Legend:

$S_{77}$  = The situation  
records for each of

${}^tT^{-1}$  = 2 dates (1901, 1929)  
for each of

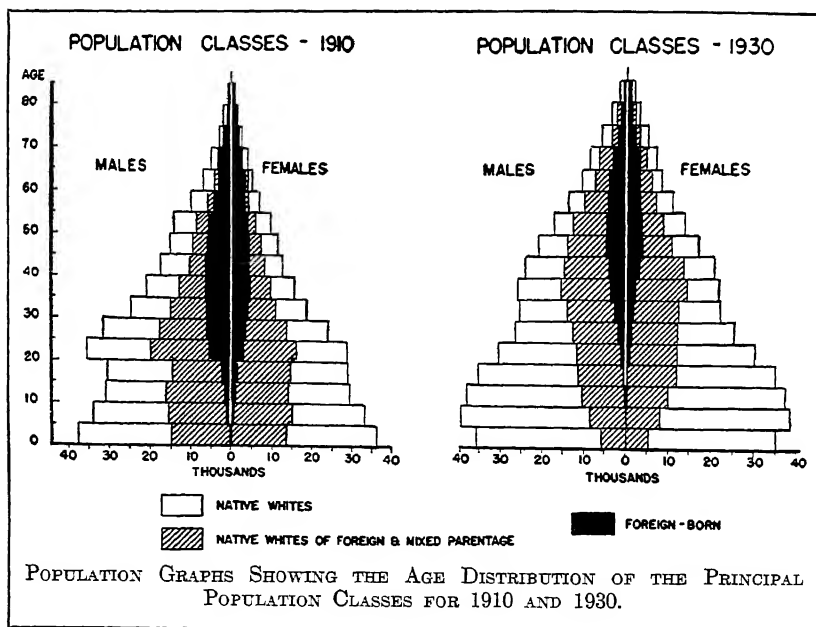
$|_p$  = the 2 sex plurels  
for every one of

${}^tT^{+1}$  = the 100 years from birth  
 $P$  = a frequency of persons surviving

*Comment:*

Age is also a change here, not necessarily a duration.  $S = T^{-2} : {}_pP$ . It is the annual change in a plurel, corresponding to calendar years. It can be treated in the usual way (as durations from variable calendar origins (birth dates) observed at annual moments). In this artificial sample the two treatments coincide as all birth dates and one calendar date have been superposed.

S. 78



Ref.: Bulletin 302, South Dakota Experiment Station, p. 24.

Descriptive formula:  $S_{78} = {}^tT^{-1} : {}_uT^{+1} : P_p : q$   
Legend:

Quantic number = 91;0;0;1

$S_{78}$  = The situation  
records for each of

$P$  = the population  
in each of

${}^t|$  = 2 dates

$|_p$  = 2 sex plurels

$T^{-1}$  = 20 years apart

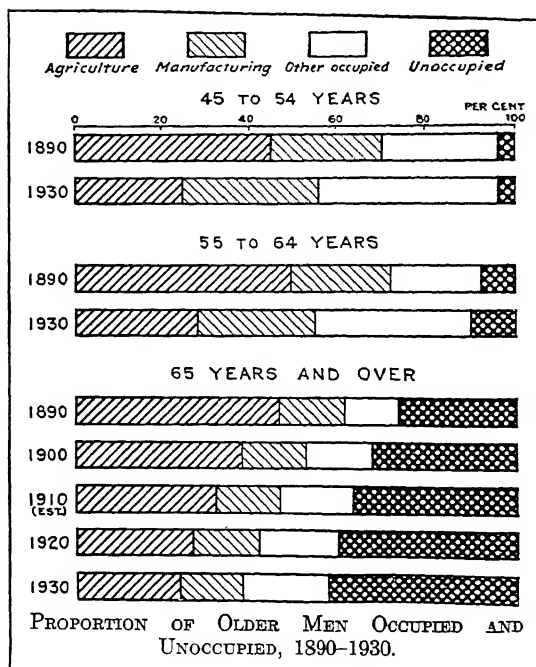
subdivided into

for each of

$|_q$  = 3 nativity plurels

${}_uT^{+1}$  = 17 5-year age classes

## S. 79



Ref.: President's Research Committee, *Recent Social Trends*.  
McGraw-Hill, Vol. I, 1933, p. 278.

Descriptive formula:  $S_{T_9} = {}^tT^{+1} : {}^tT^{-1} : {}_pP_p$       Quantic number = 19;0;0;1

Legend:

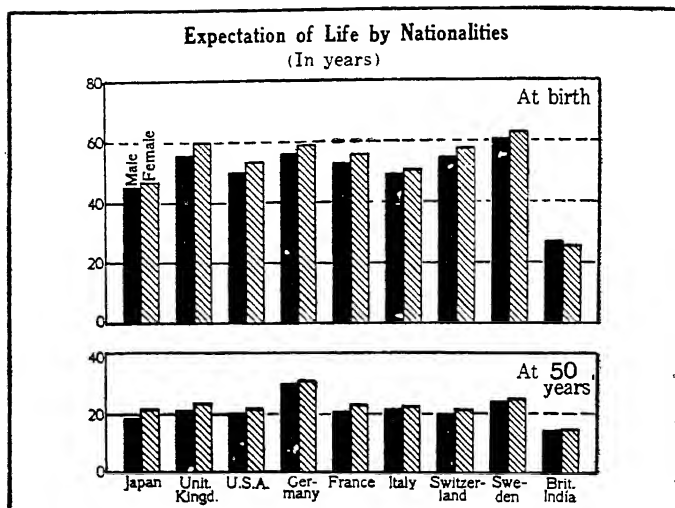
$S_{T_9}$  = The situation  
          records for each of  
 ${}^tT^{+1}$  = 3 age intervals  
          at  
 ${}^tT^{-1}$  = decennial dates  
 ${}_pP$  = the % of men  
          in each of  
 $|_p$  = 4 occupational plurels

Comment:

The two processes of societal significance are:

- (1) the adpopulating of the plurel of those retired above 65 years of age,  
+  ${}_{v\%}({}^v : P_v)$ ; and
- (2) the depopulating of the agricultural plurel at all ages, -  ${}_{v\%}P_v$ .

## S. 80



Ref.: Yano, T. and Shirasaki, I., *Nippon, A Chartered Survey of Japan*, Kikusei-Sha, Tokyo, 1936, p. 444.

Descriptive formula:  $S_{80} = {}^tT^{-1} : \underline{P}_p : q : T^{+1}$

Quantic number = 91;0;0;1

Legend:

$S_{80}$  = The situation

records at each of

${}^tT^{-1}$  = 2 ages, 0 and 50

for each of

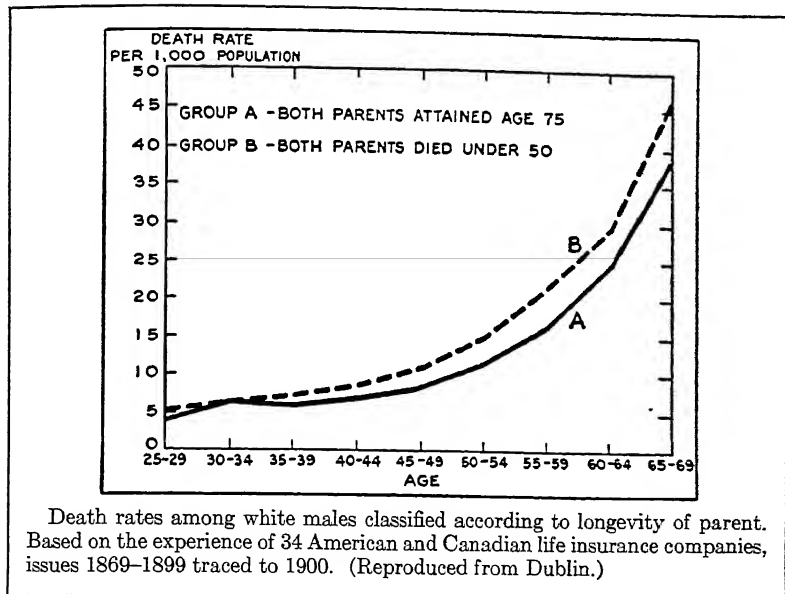
$\underline{P}_p$  = 9 national plurels

each subdivided into

$|_q$  = the 2 sex plurels

$T^{+1}$  = the expected duration of life

## S. 81



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, Vol. I, 1933, p. 617.

Descriptive formula:  $S_{81} = {}_tT^{+1} : ' : z'({}_\%PT^{-1})_p$       Quantic number = 19;0;0;1  
Legend:

$S_{81}$  = The situation  
records for each of

${}_tT^{+1}$  = 9 5-year age groups

'| = beginning at 25

the corresponding

$\%P$  = death rate

$' : z'T^{-1}$  = per year (within the limits  
1869-99)

for each of

$|_p$  = 2 plurels defined by pa-  
rental longevity

## S. 82

## MORTALITY OF INFANTS ACCORDING TO AGE OF MOTHER

Age of Mother	Austria		Norway		France
	Legitimate	Illegitimate	Legitimate	Illegitimate	
Under 17.....	2.1	4.0	2.09	4.52	6.9
17-20.....	1.7	3.0			
20-25.....	1.9	3.4	1.66	2.97	4.7
25-30.....	2.2	3.9			
35-40 } .....	2.8	4.2	2.39	4.86	4.2
40-45 } .....					4.3
45-50 } .....					6.9
50 + } .....	3.9	4.9	4.17	10.14	6.6

Ref.: Holmes, Samuel J., *The Trend of the Race*, Harcourt, Brace and Co., 1921, p. 312.

Descriptive formula:  $S_{82} = {}_tT^{+1} : ({}_pPT^{-1})_p : q$       Quantic number = 19;0;0;1

Legend:

$S_{82}$  = The situation

records for each of

${}_tT^{+1}$  = 8 age intervals of mothers

: = the corresponding

${}_pP$  = infant death rates

$T^{-1}$  = per year

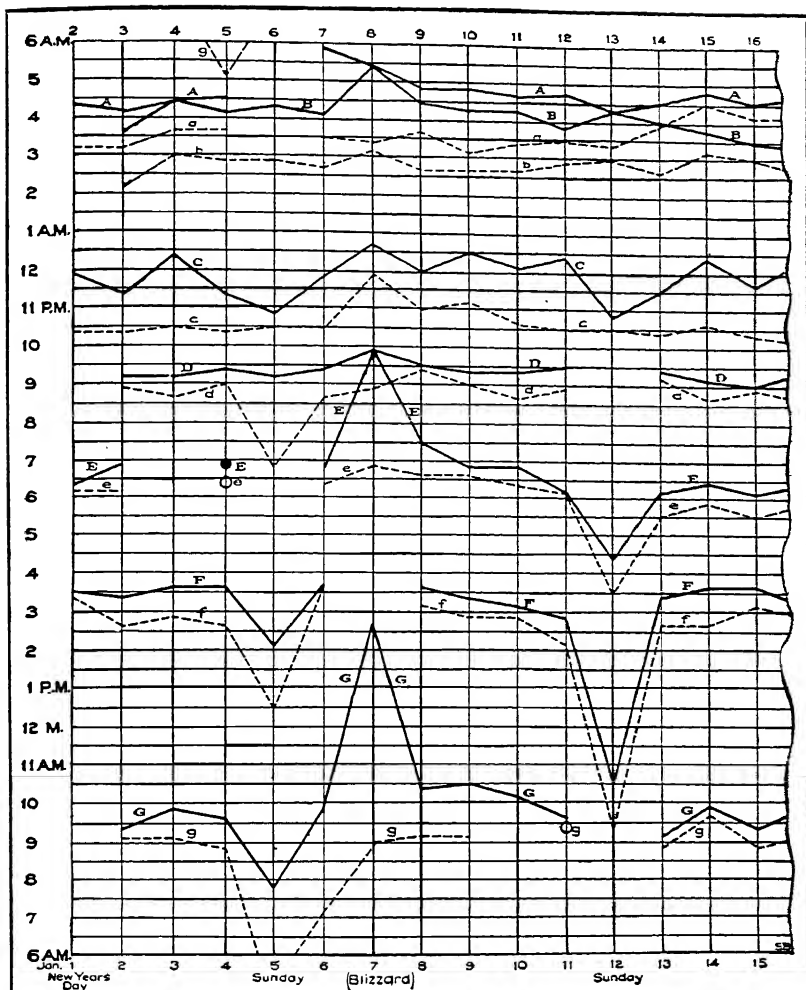
for each of

$|_p$  = 3 national plurels

: = subdivided into

$|_q$  = 2 legitimacy plurels

## S. 83



### OPERATION OF FREIGHT CAR-FLOATS AT A RAILROAD AND STEAMSHIP FREIGHT TERMINAL

Here we have time represented by days in the horizontal direction and by hours in the vertical direction. The object of the chart is to record whether car-floats are loaded and dispatched at the same hour each day.

Dotted lines show the time at which cars are pushed onto car-floats by locomotives. Solid lines show the time at which car-floats are towed away by tug-boats. Curves for any one car-float destination are in pairs bearing the same letter.

If the departure schedule is well maintained, all curve lines will be practically horizontal. Note that the blizzard of January 7 affected the locomotives less than the tug-boats.

## S. 83 (Continued)

Descriptive formula:  $S_{83} = {}^iI^0 : {}_tT^{-1} : {}^{'}T^{+1}$ 

Quantic number = 91;1;0;0

Legend:

 $S_{83}$  = The situation ${}_tT^{-1}$  = 15 days,

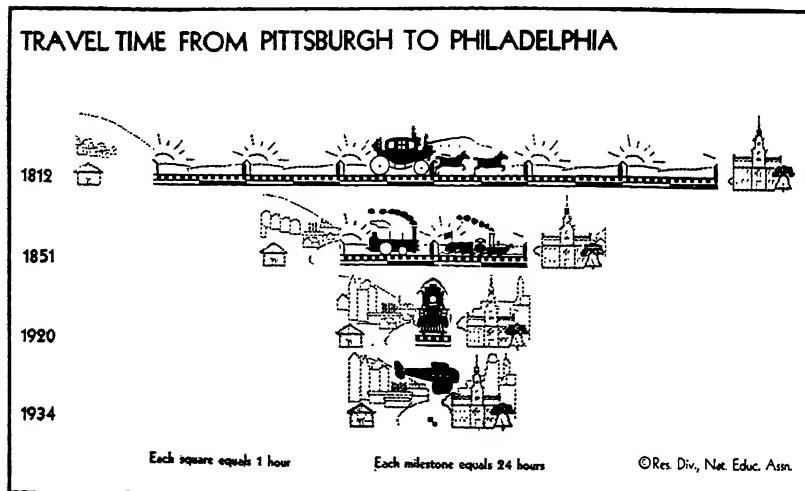
records for each of the

 ${}^{'}T^{+1}$  = 2 particular hours of departure $I^0$  = car-floats ${}^iI^0$  = 7 in number

for each of

{ loading moment  
sailing moment

## S. 84



Ref.: Research Bulletin of the National Education Association, Vol. XII, No. 5, Nov., 1934, p. 270.

Descriptive formula:  $S_{84} = {}^tT^{-1} : (T^{+1}\underline{L}^{-1})$ 

Quantic number = 91;0;9;0

Legend:

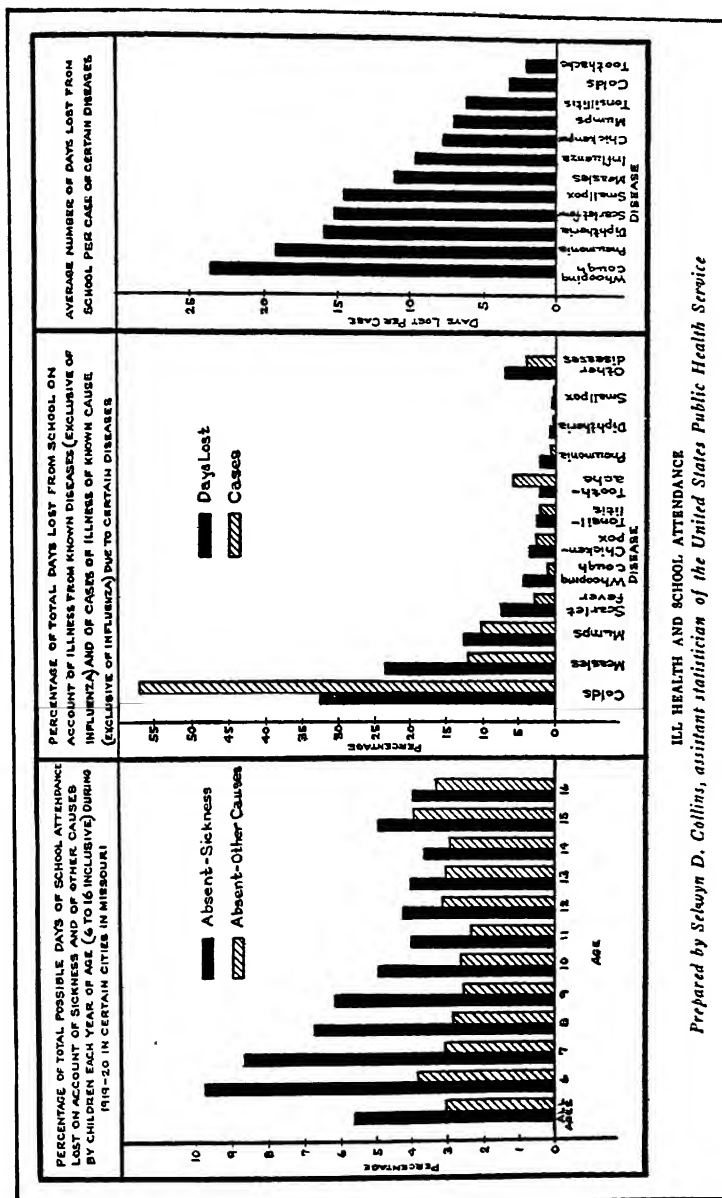
 $S_{84}$  = The situation $T^{+1}$  = the travel time

records for each of

 $\underline{L}^{-1}$  = for a particular but unmeasured distance (Pittsburgh to Philadelphia) ${}^tT^{-1}$  = 4 dates

Comment:

This situation represents a linear durationing, the reciprocal of a linear velocity.



S. 85 (*Continued*)

*Descriptive formula:*  $S_{85A} = \Sigma_{\text{t}, \text{t}} \text{T}^{+1} : \%(\text{PT}^{+1})$  *Quantic number* =  $S_{85A} = 2;0;0;1$   
*Legend:*

$S_{85A}$  = The situation (PT<sup>+1</sup>) = the pupil days lost  
 records for each of %| = in percentage units

$\Sigma_{\text{t}, \text{t}} \text{T}^{+1}$  = 11 ages and all ages

*Descriptive formula:*  $S_{85B} = \text{I}_1^0 : (\% \text{T}, \% \text{P})$  *Quantic number* =  $S_{85B} = 1;0;0;1$   
*Legend:*

$S_{85B}$  = The situation  $\% \text{T}^{+1}$  = the % of days lost,  
 records for each of and

$\text{I}_1^0$  = 12 diseases  $\% \text{P}^{+1}$  = the % of cases

*Descriptive formula:*  $S_{85C} = \text{I}_1^0 : \text{M}(\text{T}^{+1} \text{P}^{-1})$  *Quantic number* =  $S_{85C} = 1;0;0;9$   
*Legend:*

As in  $S_{85B}$  except  $\text{M}(\text{T}^{+1} \text{P}^{-1})$  = the mean days lost per case

## VI. NOTES

1. For fuller discussion of the postulates underlying the formulae for measuring and classifying societal change which are presented in this chapter, the student should read the sections on "The nature of the time concept," "Theories regarding laws of societal change," "Social versus astronomic time," etc., in Lundberg's Chapter XIII of *Foundations of Sociology*.

2. Except for Chapter IX on Durations which are represented by a vector in the time dimension.

3. For a discussion of this definition and suggestions for the scientific determination of what progress is see Ref. 14.

4. The formula for sequence is:

$$\text{tT}^{-1} = \text{a sequence} \quad (\text{Eq. 1, Ch. X})$$

This asserts an aggregation of dates,  $\text{t}$ , with indefinite intervals,  $\text{T}$ , between them. This means that the dates are in an ordinal series, i.e., their intervals are unknown.

5. S-notation is flexible in that it can express other classifications if desired. Thus Eubank's list (in quotes below) of characteristics of processes (actions) (Ref. 25, pp. 267-269) when limited to human processes, is re-expressible in part in S-notation as follows:

1. "Simple or complex":  $\text{t}(\text{I})$  or  $\text{t}(\text{I})_i$ —change of one index or of an aggregation
2. "Single or multiple" (i.e., in series):  $\text{t}, \text{T}(\text{I})$  or  $\text{t}, \text{T}(\text{I})$ —change in one period or in a series of periods
3. "Temporary or permanent":  $<, \text{T}$  or  $>, \text{T}$ —periods less than or more than a specified duration
4. "Continuous or intermittent":  $\text{t}, \text{T}$  or  $\text{t}, \text{T}$ —consecutive periods, or intermittent periods, specified by corresponding bounding dates (Appendix II, Rule 54)

5. "Functional or incidental": no direct equivalent, though correlation coefficients and time sequence express causative functionality
6. "Transitive or intransitive":  ${}_t(I) : P$  or  ${}_t(I)$ —acting on people vs. acting
7. "Reciprocal or non-reciprocal":  ${}_tI_p :: p$  or  ${}_tI_p : p$ —interaction vs. contact
8. "Singularistic or pluralistic": ' $P$  or ' $P$ —one person or many involved
9. "Approach or withdrawal or circular":  $-{}_tI_p :: p$  or  $+{}_tI_p :: p$  or  $\pm{}_tI_p :: p$ —decrease or increase, or both, of an index of social distance
10. "Opposition or accommodation":  $O_p$  or  $-{}_tD_p : p$ —"opposition" or "accommodation" (see Eq. 30 and Eq. 95, Ch. X)
11. "Similarizing or differentiating":  $-{}_t\Sigma_i$  or  $+{}_t\Sigma_i$ —"assimilizing" or "dissimilizing" (see Eq. 26, Ch. X)
12. "Integrational or disintegrational":  $+Intg$  or  $-Intg$ —"integrating" or "disintegrating" (see Eq. 102a, Ch. X)

The meaning of Eubank's terms are not always exactly the same in the S-formulae. But, it is submitted, the redefinition of these terms in S-formulae provides an equivalent set of more objective and measurable terms which are, therefore, better suited for use in a science.

6. As an illustration of the deductive derivation of any process from the S-theory equation, consider the following deductive derivation of sociating:

$$S = {}_s(T : I : L : P) {}_s \quad (\text{Eq. 7a, Ch. X})$$

To isolate populational processes, suppress I and L by considering the case where I and L have zero exponents and zero as descripts, and hence are nul and do not affect the situation, which now simplifies to:

$$S = {}_s(T : P) {}_s \quad (\text{Eq. 7b, Ch. X})$$

Since a process is being considered, the exponent of T is minus one, and the T must have a period script. Consider further only the case where the exponent of P is zero and where the class script is plural, denoting more than one qualitative plurel. Finally count the plurels, i.e., sum them up, to express their number rather than their itemized classes. The situation now becomes:

$$S = {}_tT^{-1} : P_{\Sigma p}^0 (= {}_t\Sigma_p \text{ in Brief-S}) \quad (\text{Eq. 7c, Ch. X})$$

which asserts a situation in which the number of plurels is changing. But this formula describes what is denoted by the terms effective "associating" and "dissociating," whenever the plurels are groups.

The limits of the sociating process may be noted. Letting, as usual, ' $v$ ' denote an initial date of observing the process, and ' $v'$ ', a terminal date, then

$${}''v|_{\Sigma p} \doteq 1 = \text{limit of associating, the number of plurels approaches 1 as a limit at the end of the period} \quad (\text{Eq. 8a, Ch. X})$$

$${}''v|_{\Sigma p} \doteq {}'v|_{\Sigma p} = \text{boundary between associating and dissociating, the number of plurels is the same at the beginning and end of the period} \quad (\text{Eq. 8b, Ch. X})$$

$${}''v|_{\Sigma p} \doteq {}''P = \text{limit of dissociating, the number of plurels approaches the number of persons in that population at the end of the period} \quad (\text{Eq. 8c, Ch. X})$$

The quantic number of a sociating situation is  $9;0;0;1$ , not  $9;0;0;0$ , since when more than one of the plurel-units of population  $(P^0)_p$  is taken, the populational exponent ceases to be 0 and becomes 1. Thus, in summing or counting the number of plurels,  ${}_i(P^0)_{\Sigma p} = {}_i\Sigma p$ , more than one point on the population dimension is taken and this makes an extension on it.

7. The process is the same whether the population is in one or many plurels. In the latter case it is an aggregative populating,

$$\pm {}_i P_p = \text{aggregative populating} \quad (\text{Eq. 9d, Ch. X}).$$

This dynamic aspect of population together with the static aspects constitutes larithmics, the study of populational sizes, irrespective of the quality of the people.

8. This Brief-S formula may be written as the full formula and for more than one interrelation and period, either as:

$$\begin{aligned} {}_i T^{-1} : {}^p P_p :: P_p : (I)_i &= \text{change of static interrelations} & (\text{Eq. 12b, Ch. X}) \\ \text{or } {}^p P_p :: {}^p P_p : {}_i (IT^{-1})_i &= \text{dynamic interrelations, interacting proper} & (\text{Eq. 12c, Ch. X}) \end{aligned}$$

The quantic number is the same for both, namely,  $9;j;1;2$ .

9. Thus processes of "isolating" and "contacting" may be distinguished. Isolating is measurable as a decrease in the isolation index. (Eq. 11, Ch. VII.) Contacting might be measured either by some index summarizing the contact matrix (Eq. 12, Ch. VII) when its cell entries are dynamic components or, by the change in some index which summarizes the contact matrix when built up of static components. (See S. 54 and 55, Ch. X.)

10. The derivation of these formulae is as follows:

For the first plurel,  $P_1$ , the sum of its losses in one period is given by the total of the first row of cells  $(-{}_i P_1 : \Sigma p)$  in Eq. 13, Ch. X, and the sum of its gains is given by the total of the first column of cells  $(+{}_i P_{\Sigma p} : 1)$ . The algebraic sum of gains and losses is the net mobility for that plurel:

$$(-{}_i P_1 : \Sigma p) + (+{}_i P_{\Sigma p} : 1) = (\pm {}_i P_1) = \text{net mobility of one plurel} \quad (\text{Eq. 15a, Ch. X})$$

while the absolute sum of gains and losses is the gross mobility for that plurel:

$$| -{}_i P_1 : \Sigma p | + | +{}_i P_{\Sigma p} : 1 | = | {}_i P_1 | = \text{gross mobility of one plurel} \quad (\text{Eq. 15b, Ch. X})$$

Similarly for the plurels of an aggregation the net and gross mobilities are  $(\pm {}_i P)_p$  and  ${}_i P_p$  respectively.

Next, to derive the index of mobility in an aggregation of plurels, first divide the net mobility of each plurel in the aggregation of  $p$  plurels by the total population in the aggregation, and multiply by 100 to express the net mobilities in percentage terms:

$$100(\pm {}_i P)_p P^{-1} = (\pm \% {}_i P)_p \quad (\text{Eq. 15c, Ch. X})$$

The algebraic sum (and therefore the mean) of these net mobilities of the plurels is zero by definition, as every gain to any plurel is an equivalent loss to

some other plurel. (If it is not zero, the populating process is also going on and its amount must be subtracted to isolate the mobility process.)

$$\sum_1^h (\pm \pi_{it} \cdot P) = (\pm \pi_{it} \cdot P) \Sigma_p = 0 \quad \text{and} \quad \mathbf{M}(\pm \pi_{it} \cdot P_p) = (\pm \pi_{it} \cdot P) \Sigma_p P^{-1} = 0$$

(Eqs. 15d, e, Ch. X)

Thus each gain or loss is a deviation from the mean. The standard deviation of these gains and losses is the root mean square:

$$((\pi_{it} \cdot P_{\Sigma p}^2) p^{-1})^{.5} = \sigma \quad (\text{Eq. 15f, Ch. X})$$

For interpreting this  $\sigma$  it is more intelligible if expressed as a percentage of the maximal  $\sigma$ , so that the resulting index of mobility can be thought of as a percent of maximal possible mobility and compared in different situations. The maximal sigma occurs when one of two groups has all the population at first and loses them all, so that the other group gains them all. Thus one deviation is  $-100\%$ , another is  $+100\%$ , and all the others are  $0\%$ , or, expressed in generalized algebraic terms:

$$(\pi_{it} \cdot P_{\Sigma p}^2 p^{-1})^{.5} = (2(100\%)^2 p^{-1})^{.5} = 100\% (2/p)^{.5} = \sigma' = \text{maximal sigma}$$

(Eq. 15g, Ch. X)

Dividing Eq. 15f by Eq. 15g gives:

$$\sigma/\sigma' = (\pi_{it} \cdot P_{\Sigma p}^2)^{.5} / 100 \times 2^{.5} \quad (\text{Eq. 15h, Ch. X})$$

This is a proportion of unity and has to be multiplied by 100 to convert it to percentage units giving:

$$100\sigma/\sigma' = (.5 \pi_{it} \cdot P_{\Sigma p}^2)^{.5} = \text{Mb}_N \quad (\text{Eq. 14c, Ch. X})$$

This is a percentage measure in sigma terms of the amount of rearranging of the sizes of the plurels within the total population studied.

For the derivation of the formula for gross mobility, since this has no natural maximum as net mobility has, the conventional "100% turnover" may be taken as a standard for deriving a percentage measure. This leads to adopting the simple sigma (Eq. 15f) of percentage gains and losses as the index of mobility. Only, since every shift of membership is counted twice, once as a gain to some plurel, and again as a loss to some plurel, it has to be divided by 2 so that, although 100% of the membership is lost to a plurel and replaced by another 100% of members, it will be called a 100% turnover and not a 200% turnover. By inserting 2 in the denominator of Eq. 15f there results Eq. 14b, the index of gross mobility. Note that gross mobility is a function of an average for the  $p$  plurels, while net mobility is a function of an average for 2 plurels, which is the situation of maximal net mobility.

These two indices of mobility are sigmas (in percentage form) of the main diagonal cell entries of the mobility matrix (Eq. 13, Ch. X). Other indices measuring mobility can be derived, based on all the cells of the matrix, but usually the data in this form are not obtainable. Firms keep complete records of their hirings and firings, but not ordinarily of the other firms from which or

to which their employees transfer. Similarly for other types of plurels, complete records for filling in all cells of the matrix are usually lacking.

The special case of mobility among plurels whose membership overlaps is more complex. The matrix (Eq. 7, Ch. VII) states in the cells the number of persons belonging to every pair of plurels, as in multi-membership in fraternal lodges. In order to *completely* represent mobility here, one technic is to construct a matrix like Eq. 13, Ch. X, but enlarged to provide a row and column for every pair (or larger combination if desired) of plurels, i.e., an array for every cell of Eq. 7, Ch. VII. (Note that as this matrix Eq. 7, Ch. VII is symmetric, there are only  $\frac{p^2 + p}{2}$  different cell entries.)

This enlarged matrix would show the effects of the transfer of a person upon all the plurels of which he was a member. Thus, suppose a person belonging to plurels  $P_1$  and  $P_2$  resigned from plurel  $P_2$ , and joined plurel  $P_3$ . The cell in the row of the enlarged matrix representing the  $P_{1,2}$  combination of groups and in the column representing the  $P_{2,3}$  combination would have a 1 added to its cell-entry, which can be symbolized by  $P_{(12)(13)}$ .

Other technics for dealing with multiple-membership situations can readily be devised by the aid of matrices to insure complete and systematic treatment.

The standard deviation is usually preferable to the average deviation for statistical calculations, and is here made the basis of these mobility indices. If, however, the average deviation is used, the indices of mobility simplify to:

$${}_A\text{Mb}_N = |(\pm \%v \cdot P)| \Sigma p / 2 = \text{index of net mobility in average deviation units} \\ (\text{Eq. 16a, Ch. X})$$

This is simply the percentage of transfers in the whole population, and

$${}_A\text{Mb}_G = |\%v \cdot P| \Sigma p / 2p = .5\%v \cdot P_p = \text{index of gross mobility in A.D. units} \\ (\text{Eq. 16b, Ch. X})$$

This is simply the average percentage of turnover among the plurels of the whole population.

Eq. 16 is readily derived by substituting the A.D. for  $\sigma$  in Eq. 14. As before, this index of gross mobility in A.D. units is zero only when no person changes plurels, and rises to 100% when every person changes plurels once on the average, and rises to 200% when everyone changes twice on the average, etc. Similarly, the index  ${}_A\text{Mb}_N$  is zero either when no changes occur, or when gains equal losses in every group, so that no net changes of size occur; and it rises to a maximum of 100% when one plurel which initially had members loses them all to a plurel which initially had no members.

11. As an example of the use of these mobility formulae, consider Sorokin's propositions (Ref. 68, p. 160):

1. "The principal forms of social mobility of individuals and social objects are: horizontal and vertical. Vertical mobility exists in the form of ascending and descending currents. Both have two varieties: individual infiltration and collective ascent or descent of the whole group within the system of other groups.

2. "According to the degree of the circulation, it is possible to discriminate between immobile and mobile types of society.

3. "There scarcely has existed a society whose strata were absolutely closed.

4. "There scarcely has existed a society where vertical mobility was absolutely free from obstacles.

5. "The intensiveness and the generality of vertical mobility vary from group to group, from time to time (fluctuation in space and in time). In the history of a social body there is a rhythm of comparatively immobile and mobile periods.

6. "In these fluctuations there seems to be no perpetual trend toward either an increase or decrease of vertical mobility.

7. "Though the so-called democratic societies are often more mobile than autocratic ones, nevertheless, the rule is not general and has many exceptions."

Asserted in S-notation in quantitative forms susceptible of more exact determination, these are:

$$Mb = (.5\%v.P_{2p}^2)^{.5} \quad (\text{Eq. 14, Ch. X})$$

The size of Mb measures Proposition #2 above in the respects defined by the plurals, p. By defining p as either "vertical" or "horizontal" groups (see Chapter VII for discussion of these terms) and by recognizing that P, the persons, may have  $|_q$  subplurals substituted, Proposition #1 above is expressed:

$$(0 < Mb < 100\%)_{\%2p} \quad (\text{Eq. 17, Ch. X})$$

This assertion that mobility tends to be greater than zero and less than complete expresses Propositions #3 and 4. The exact degree of probability with which mobilities less than 100% and more than zero occur is expressed by the class script as a percent of all the populations studied,  $|\%2p$ . This can measure exactly the phrases "tends to be" or "scarcely" in Sorokin's wording. Eq. 17 is verbalized as that in a certain percent of all the plurals mobility is greater than zero and less than 100%.

$\sigma(Mb_p) > 0$ , the sigma of the index of mobility in  $|_p$  populations is not zero  
(Eq. 18, Ch. X)

$\sigma(^tMb) > 0$ , the sigma of the index of mobility among  $^t|$  dates is not zero, i.e., varies  
(Eq. 19, Ch. X)

These measure Proposition #5 as to generality, while the time period specified in calculating Mb determines the velocity which is taken as the essence of "intensiveness."

$Mb \cdot T = 0$ , zero correlation of mobility with time, i.e., no trend  
(Eq. 20, Ch. X)

This asserts zero correlation (a scalar product equal to zero) between mobility and time, i.e., the absence of any trend for mobility to either increase or decrease with time. (This is an instance of primitive "recovarying" (see under Indicatory processes below) where one of the two indices is the implicit attribute-time product ( $I^oT$ ).)

$$Mb \cdot I_D > 0 \quad (\text{Eq. 21, Ch. X})$$

This asserts a positive correlation between mobility and an index of democracy-autocracy. To convert this to an equality of specified amount is the challenge to further research.

12. In Brief-S the cross script  $P_1$  denotes a population of some aggregation of geographic spaces, so that the formula for net migrating is :

$$(.5(\%_t P^2)_{\Sigma 1})^{.5} = Mg_N = \text{net migrating} \quad (\text{Eq. 22a, Ch. X})$$

Gross migrating may be similarly formulated from Eq. 14b as :

$$(.5l^{-1}(\%_t P^2)_{\Sigma 1})^{.5} = Mg_G \quad (\text{Eq. 22b, Ch. X})$$

13. The full descriptive formula for the economic process is a fourth-degree matrix equation (distributed in three sectors—T, I, P) as follows :

$$S = P_p :: P_P : (IT^{-1})_m = \text{the economic process in full} \quad (\text{Eq. 23b, Ch. X})$$

The number of dimensions in this formula is determined by the sum of the four class scripts,  $|\Sigma_p| + |\Sigma_P| + |\Sigma_m| + 1 = n$ . The  $t$  class-intervals of time are all on one time dimension,  $|t|$ . The total "order" of the matrix, which is also the number of its cells, is given by the product of the descripts  $|\Sigma_p| \times |\Sigma_P| \times |\Sigma_t| \times |\Sigma_m|$ . (Note that each descript is summed to denote a single number, since, when unsummed, it represents an aggregation.) This is at maximum an enormous number, for the number of parties,  $|\Sigma_p|$ , can be anything up to the total number of living persons (about two billion), plus estates of dead people, plus all groupings of people, such as families, partnerships, companies, associations, governments, etc., with all their departments and subdivisions which act as entities in buying and selling goods or services. For a wild guess, suppose there are a billion such economic groups ( $\pm$  half a billion, perhaps, as a guess at the error of estimate of the main guess). This makes three billion times three billion as the maximum order of the second-degree matrix, Eq. 23, Ch. X, or  $9 \times 10^{18}$ , as the number of its cells,  $|\Sigma_p| \times |\Sigma_P|$ . Such astronomical quantities can be handled in our matrix notation quite as easily as simple matrices of a half dozen buyers and sellers. Matrix algebra provides a generalized scheme for manipulating these quantities in rigorous ways, regardless of their arithmetic size. As a practical fact, however, this maximum is purely theoretic, since most of the three billion (?) parties make no exchanges with each other and show zero cell entries in the matrix. Actually each party will interact with a very limited number of other parties, usually local ones. The range of the number of interactants exchanging commodities with any one party runs from 1 (as when a person sells all his services, paid for in a salary, to one party) to perhaps one or two hundred million at most (as when people buy stamps of a government).

The vast majority of the parties from whom a person buys or to whom he sells is perhaps of the order of a few dozen, when the rural masses of Asia, etc., are remembered as averaging down the many interactants of each party in a metropolis.

The number of periods is any number for which the investigator gathers data—12 months, 1 year, or a century, or other time series. In currently published tables and graphs of time series,  $|\Sigma_t|$  averages perhaps around the order of a dozen or two.

The number of commodities bought and sold is, again, enormous, depending on how finely the investigator classifies those commodities. In Eq. 23a, Ch. X the least classifying possible is represented, for here only one class is portrayed, namely, anything exchanged for money (in the widest sense). From this lower limit it may go on, up to the inventory numbers of the largest mail order firms, and summed for all kinds of firms and other exchangers, perhaps reaching millions or even billions, if services of persons are individualized far enough.

14. A further trend towards directional flow within the matrix can be set up by arranging so that prime producers (farmers, miners, fishermen, etc.) are in a section together, processors and manufacturers are sectioned together, and the residual, non-gainfully employed parties are in a pure consumers' section together. Let the matrix be represented not on a plane, but as the surface of a cylinder, with producers to consumers arranged in clockwise order. Then money will tend to flow in a counter-clockwise direction, and commodities will tend to flow in the reverse clockwise direction. Of course, this will be only a tendency, since many commodities are absorbed by processors and distributors, as these are also ultimate consumers. This tendency, however, for a producer-consumer directional flow lends itself to many types of economic analysis.

15. A variant of this is the "financial statement" as of some date. This is a cumulation of profit and loss up to that point in time. The assets are cumulated income; the liabilities are future expenditures already legally incurred; and the difference is the proprietorship, which is equivalent to a net accumulated profit and loss.

16. Economics was selected for simplicity of illustration. The interacting matrix, however, is just as applicable to political, religious, interracial, or any other field of interhuman behavior. In Economics the indicators are more often quantitative ones,  $I^{+1}$ , while they are more often only qualitative,  $I^0$ , in other social sciences. Also in Economics the quantitative indicators can more often be expressed in a common (monetary) unit than is the case in other social disciplines. Wherever objective indicators, qualitative or quantitative, can be developed for any societal phenomena of interaction, the matrix form of analysis can be applied and S-theory notation can be used as a tool.

17. Two subforms of differing may be distinguished, namely, to change a qualitative entity into some other qualitative entity, as illustrated above, and a continuing qualitative activity such as, sweeping, sewing, walking, or speaking. In the full descriptive formulae these are distinguished as change of a static attribute and as a dynamic attribute:

$$\begin{aligned} {}^tT^{-1} : I^0 &= \text{differing proper, a changing quality on successive dates} && \text{(Eq. 25b, Ch. X)} \\ {}_t(I^0T^{-1}) &= \text{"acting," a qualitative activity} && \text{(Eq. 25c, Ch. X)} \\ &&& \text{(See S. 16, 17, Ch. X)} \end{aligned}$$

For a hierarchy of attributes the formula is:

$$I_1^0 : I_2^0 : \dots : I_n^0 = {}_tI_n^0 = \text{"hierarchical differing," a changing system, or hierarchical pattern, of qualities} \quad \text{(Eq. 25d, Ch. X)}$$

For an aggregate of changing attributes the formula is:

$$({}_t\text{I}^0)_1 \quad (\text{Eq. 25e, Ch. X})$$

while for a change in the composition of an aggregate of attributes, it is:

$$({}_t\text{I}^0) \quad (\text{Eq. 25f, Ch. X})$$

Eq. 25 merely states the fact of a distinguishable qualitative change. Whenever qualities can be quantified, the degree of similarity between two qualities may be measurable by a correlation coefficient. Interpreting this as a cosine of an angle measures degrees of qualitative change in terms of degree of angle, as described in Chapter VI. For examples of "differing" see S. 19, 20, 21, and 22, Ch. X.

18. The limit of assimilizing in unity, the boundary between assimilizing and dissimilarizing is the initial number of qualities, and there is no definite limit of dissimilarizing as there is for dissociating. (Cf. Eq. 8, Ch. X.)

$$\begin{array}{ccccccc} 1 & \leftarrow & \frac{-}{t}|\Sigma_i & \leftarrow & \frac{''}{t}|\Sigma_i & \rightarrow & \frac{+}{t}|\Sigma_i \rightarrow \infty \\ \text{Lower} & & \text{Assimi-} & & \text{Bound-} & & \text{Dissimi-} \\ \text{limit} & & \text{larizing} & & \text{ary} & & \text{larizing} \\ & & & & & & \text{Upper} \\ & & & & & & \text{limit} \end{array} \quad (\text{Eq. 26a, Ch. X})$$

19. Since indicators are either attributes or indicants, the term "indicatory change" may conveniently denote changing indicators:

$$({}_t\text{I}_1^0) = \text{"indicatory change"} \quad (\text{Eq. 27, Ch. X})$$

A changing indicant is symbolized as:

$$\begin{aligned} \pm {}_t\text{I}^1 &= \text{indicating} & (\text{Eq. 28, Ch. X}) \\ &(\text{See S. 29, 30, 31, 39, and 43, Ch. X}) \end{aligned}$$

"Increasing" and "decreasing" are the terms used here with complete generality for processes in any sector, or combination of sectors.

20. "Intentional" has the meaning of "purposeful," but can be behavioristically defined by the test of whether the intended result is *announced beforehand*.

21. Grading is from the Latin root meaning "a step," "progress" being a step forward, "regress," a step backward. Grading connotes in English the classifying of some quality, or value, into degrees, or steps.

The two subtypes of grading are symbolized by:

$$\begin{aligned} \pm {}_t\text{V}_{\text{Cu}} &= \text{"cumulating"} & (\text{Eq. 32, Ch. X}) \\ &= \text{change of the amount of a value without human effort} \\ + {}_t\text{V}_{\text{Cu}} &= \text{"accumulating"} & (\text{Eq. 32a, Ch. X}) \\ - {}_t\text{V}_{\text{Cu}} &= \text{"decumulating"} & (\text{Eq. 32b, Ch. X}) \\ \pm {}_t\text{V}_{\text{Co}} &= \text{"operating"} & (\text{Eq. 33, Ch. X}) \\ &= \text{change of the amount of a value caused by human effort} \\ + {}_t\text{V}_{\text{Co}} &= \text{"co-operating"} & (\text{Eq. 33a, Ch. X}) \\ - {}_t\text{V}_{\text{Co}} &= \text{"maloperating"} \text{ (i.e., destroying or wasting a value, or creating a} \\ &\quad \text{danger, a disease, a social problem)} & (\text{Eq. 33b, Ch. X}) \end{aligned}$$

22. The tensing process might be classed as a single complex process, since it is an index combining populational and indicatory indices and might be discussed later in this chapter in the section on momentum. As the P cancels out

reducing E to a ratio of two indicants, it also can be classed (as it is here) as a single, simple, primary, indicatory process. It also fits here as part of the discussion of the tension theory following the processes of valuating and co-operating, as defined by that theory.

23. Ordinarily attensing goes on when recruiting, evaluating, and regressing are combined, though sometimes large amounts of one of these processes may offset change in the opposite direction in others, since the process of attensing is defined as a positive change in the product inside the parenthesis of Eq. 35 above.

$$+_t(PDV^{-1}) = +_tE = \text{definition of attensing} \quad (\text{Eq. 35a, Ch. X})$$

so that:

$$[(+_tP_v) + (+_tD_v) + (-_tV)] \cdot (+_tE_v) = r > 0 \quad (\text{Eq. 36a, Ch. X})$$

recruiting + evaluating + maloperating are positively correlated with attensing and:

$$[(-_tP_v) + (-_tD_v) + (+_tV)] \cdot (-_tE_v) = r < 0 \quad (\text{Eq. 36b, Ch. X})$$

conflicting + accommodating + co-operating are positively correlated with detensing

Decreases in the numerator of Eq. 35 and increases in the denominator will always decrease the tension ratio producing detensing:

$$-_t(PDV^{-1}) = -_tE = \text{definition of "detensing"} \quad (\text{Eq. 35b, Ch. X})$$

24. For another sociological analysis see Ref. 39.

25. A more detailed comparison of Von Wiese's formula for a process and tensing follows:

In S-notation where objectively observable indicators (data) represent the characteristics (phenomena) which he symbolizes directly, a process is a resultant change in some index over a series of dates  ${}^u(I)$ ; an attitude is designated by some index  ${}^t(I)_s$  or indices at some antecedent series of moments; and the situation is summarized in another antecedent index or combination of them  ${}^t(I)_s$ .

The combination is suggested, pending further research, as a hierarchical aggregation of situational indices and is so denoted by using the plural descript, the small letter s.  $|_s = |_{1:j:\dots:z}$ . In the simplest instance an antecedent attitude and a situation are taken each for one date only,  ${}^t|$ . But for completer study attitudes and situations may be observed on a series of dates,  ${}^t|$ , and correlated as a time series (allowing for possible lag) with the process index,  ${}^u(I)$ . Then Von Wiese's formula for a social process, in S-notation, is:

$${}^t(I)_s(I)_s = {}^u(I) \quad (\text{Eq. 38b, Ch. X})$$

As every equation connotes perfect correlation between its left-hand and its right-hand members as these may vary, this can be recast into a correlational equality in order to make it comparable with the imperfect equalities, or tend-

encies towards equality, exhibited by most correlated variables in the social sciences. In this form, the scalar product (correlation coefficient,  $r$ ) of the index of the process  ${}^u(I)$  and the arithmetic product of the indices of attitude  $(I)_s$  and of the situation  $(I)_s$  are unity.

$${}^t((I)_s(I)_s) \cdot {}^u(I) = r = 1.00 \text{ in S-notation (Eq. 40a, Ch. X)}$$

or more roughly with less specification of the symbols and the attendant conditions:

$$AS \cdot P = 1.00 \text{ in Von Wiese's notation (Eq. 40b, Ch. X)}$$

This is now in parallel form to the tension hypothesis, Eq. 39a, Ch. X. In comparing the two, the important difference is that Von Wiese's formulation of the causal factors of a societal process is in terms of ideal abstractions yielding an assumed but unverifiable perfect correlation; the tension theory formulation is in less inclusive terms of measurable entities yielding some verifiable degree of correlation. For science verifiable hypotheses are to be preferred to unverifiable ones. As always, when the correlation,  $r$ , has been observed, the coefficient of non-determination,  $k^2(k^2 = 1 - r^2)$ , states the percentage of unknown or unmeasured causal influences which constitute the residual causation of the societal process,  ${}^u(I)$ , in question. The percentage-coefficient of determination,  $100r^2$ , states the percentage degree to which the tension hypothesis, as measured in any given situation, is verified. The problem for research then is to discover whether, as the desires and the desiderata are measured with increased reliability, validity, and completeness under specified conditions, the correlation increases between the measured tension and the societal process which is its alleged effect.

A further but minor point of difference between the tension hypothesis, Eq. 39a, and Von Wiese's hypothesis, Eq. 40, is that the former involves a ratio of the causal factors,  $E = PD/V$ , while the latter involves their product,  $AS$ . Von Wiese disclaims mathematical rigor in his formula and probably suggested a "product" as meaning only some combination in which each influenced the other in giving the result. The tension theory assumes a ratio since, as in the economic ratio of demand and supply, the societal tension is observed to vary directly as the attitudes of desire and inversely as the amount of the desideratum in the situation. In the special but very frequent case of a qualitative desideratum,  $V^0$ , the ratio and the product become equivalent, since dividing by 1 is arithmetically equivalent to multiplying by 1.

26. For uniform nomenclature secondary processes are christened with terms involving the prefix "re-" just as nullary process terms end in "-izing," and primary process names end in "-ating." (There are a few exceptions in order to avoid forcing terms.) Among secondary processes the syllable "co-" distinguishes the recorrelating from the redeviating processes.

27. Note that in the subjoined table (i.e., matrix) of second-order moments (taken about the mean) the variances are the main diagonal entries and the covariances are the other entries.

	I	J	K		Z
I	$\frac{\Sigma I^2}{P}$	$\frac{\Sigma IJ}{P}$	$\frac{\Sigma IK}{P}$		$\frac{\Sigma IZ}{P}$
J	$\frac{\Sigma JI}{P}$	$\frac{\Sigma J^2}{P}$	$\frac{\Sigma JK}{P}$		$\frac{\Sigma JZ}{P}$
K	$\frac{\Sigma KI}{P}$	$\frac{\Sigma KJ}{P}$	$\frac{\Sigma K^2}{P}$		$\frac{\Sigma KZ}{P}$
Z	$\frac{\Sigma ZI}{P}$	$\frac{\Sigma ZJ}{P}$	$\frac{\Sigma ZK}{P}$		$\frac{\Sigma Z^2}{P}$

$= (I)_{i,i}$  = matrix of second moments of  $i$  indices (each about its mean)  
(Eq. 43a, Ch. X)

If the moments are expressed in standard deviation units in order to reduce them all to expression in comparable units, the matrix becomes the familiar table of intercorrelation coefficients with unities in the main diagonal cells, since

$$\frac{\Sigma IJ}{P\sigma_I\sigma_J} = r_{IJ} \quad \text{and} \quad \frac{\Sigma I^2}{P\sigma_I^2} = \frac{(\sigma_I)^2}{(\sigma_I)^2} = 1 \quad (\text{Eqs. 44a and b, Ch. X})$$

	I	J	K		Z
I	1	$r_{IJ}$	$r_{IK}$		$r_{IZ}$
J	$r_{JI}$	1	$r_{JK}$		$r_{JZ}$
K	$r_{KI}$	$r_{KJ}$	1		$r_{KZ}$
Z	$r_{ZI}$	$r_{ZJ}$	$r_{ZK}$		1

$= (rI)_{i,i}$  = matrix of intercorrelations of  $i$  indices  
(Eq. 43 b. Ch. X)

The variance is thus seen to be but a special case of the covariance which when expressed in sigma units (reducing it to a 0 to 1.00 scale) is the correlation coefficient. The kinship of redeviating and recorrelating processes may be thus clarified. (Eq. 41 can be written " $I_{I,I}$  = redeviating" to show its parallelism to Eq. 42, but the sigma notation is more familiar and convenient than the scalar self-product notation in this case.) In terms of vectorial algebra, Eq. 43b is a matrix of scalar products of *unit* vectors. The unit is sigma as shown by writing it in the class-interval script within the parenthesis of this index in Eq. 43b.

28. A special subform of this process is "discriminating," defined by Hankins as unequal treatment of equals (Ref. 28). In our notation this is dispersing in a population whose sigma has been, or theoretically should be, zero. Thus if members of one sect or race in a country whose constitution declares all races and religions equal before the law are treated differently from other sects or races, discriminating is going on in respect to those indices of differential treat-

ment. Thus, negro school children, if they receive less per capita of State school funds than white children in the United States, are discriminated against.

29. Concepts of Physics are similarly defined. Thus a "velocity" is a ratio of distance to time, regardless of whether the "content" is the velocity of a bullet, a feather, or a train, and regardless of whether it is expressed in units of feet per minute, centimeters per second, or otherwise.

30. See S. 36, 39, 47, Ch. X for situations in which by calculating this sigma the reordinating process would be measured. The dispersion of dynamic indices, such as annual income, is an example of annual reordinating.

31. Redispersing and reordinating are related as a difference in two sigmas on two dates versus a sigma of the differences between two dates.

$$v''(\sigma I) - v'(\sigma I) = \pm v'(\sigma I) = \text{redispersing} \quad (\text{Eq. 45c, Ch. X})$$

$$\sigma(I) = \sigma(v'I - v'I) = v''(\sigma I)^2 + v'(\sigma I)^2 - 2 v'' \cdot v'(I) v''(\sigma I) v'(\sigma I) \quad (\text{Eq. 46b, Ch. X})$$

(Cf. Eq. 18b, Ch. V)

Thus the dispersion of incomes on date  $v'$  subtracted from the dispersion of incomes on date  $v''$  shows the net change in dispersion. But if the changes of income of the persons are calculated first,  $v'(I)$ , and the sigma of these then calculated, the reordinating of the incomes is shown, for this includes both change of dispersion and also reranking of the persons in respect to income. The correlation coefficient,  $v'' \cdot v'(I)$  ( $= r_{12}$  in the conventional but less specific notation) may be less than unity, measuring reranking of the persons. Reordinating can thus be calculated directly as the sigma of differences, or, indirectly by compounding it out of the processes of redispersing and reranking (see below) by formula, Eq. 46b, Ch. X.

32. The velocity of this process of competing was 1% per year. This is an average velocity, disregarding fluctuations from the steady trend for foreign shipping to out-compete American vessels. (See S. 9, Ch. XI.)

33. The derivation and a special case or two may be noted here. The algebraic sum of the changes is zero, i.e.,  $\Sigma_v V/P = 0$ , as gains must balance losses by definition. The case of competitors who are persons (not plurels) is followed in the notation below. By suitable shift of notation the formulae cover both.

The standard deviation of these changes is given by:

$$\sigma_{v'}^2 = \Sigma_v V^2/P \quad (\text{Eq. 48, Ch. X})$$

This is maximal when one loses all the value and another gains it all, while others have none. Thus two cases show 100% deviations, and the rest 0% deviations. This maximum  $\sigma$  is given by:

$$\max \sigma_{v'} = (2(\Sigma V)^2/P)^{.5} = 100/(\cdot 5P)^{.5} \quad (\text{Eq. 49, Ch. X})$$

The ratio of these two sigmas (i.e., the ratio derived as the square root of Eq. 48 divided by Eq. 49 multiplied by 100 to make the ratio a percentage) gives Eq. 47, when it is recalled that by definition  $\Sigma V = 100\%$ .

Consider one special case where the competitors start with nothing ("at scratch"). Note that here, if the change,  $V'$ , is given not as a difference in amounts on two dates as in the case above, but as events, earnings, or happen-

ings in a period,  $\epsilon V$  becomes each competitor's percentage share of such change expressed as a deviation from the mean share  $(100\%/P)$ .

$$\epsilon V = \frac{100V'}{\Sigma V'} - \frac{100}{P} \quad (\text{Eq. 50, Ch. X})$$

Thus four basketball teams winning points in a season of 400, 200, 100, and 300 respectively, converted into percentages and then into deviations from the mean, give an index of competing of 15.8%.

$(C_p = (.5(15^2 + 5^2 + 15^2 + 5^2))^{.5})$  Whereas, were their winnings more equal, as in 300, 230, 210, 260 points, the effective competing would be reduced to  $C_p = 4.8\%$ , as here there is less dispersion and no reranking of the competitors. In this situation, as no team plays in all games, it is physically impossible to obtain a monopoly of the points as it would be were the number of competitors reduced to two. Even with two competitors the maximum  $C_p$  is 50%, as their status at the start is equal, and only at the end is it a monopolistic status. 100% competing would be represented by the initial possession of a championship cup by one competitor who loses it while the other gains it, thus reversing one initial monopolistic status. Such a cup is an all-or-none value,  $V^0 = 1$ .

A second special case of common occurrence is when the data are given as a ranking. The formula here is:

$$C_p = 100(.5 - .5r)^{.5} = C_p \text{ from rankings} \quad (\text{Eq. 51a, Ch. X})$$

$$= 71\sqrt{1 - r}$$

For the sigma of gains or losses in ranks it is:

$$\epsilon\sigma = \epsilon'\sigma(2(1 - r))^{.5} \quad (\text{Eq. 51b, Ch. X})$$

where

$$\epsilon\sigma^2 = \Sigma \epsilon_i^2 V_{v'}^2 P^{-1} \quad \text{and} \quad \epsilon'\sigma^2 = \Sigma \epsilon_i^2 (V_{v'}^2) P^{-1}$$

Eq. 51b is the usual sigma of a difference when the two sigmas are equal, as the initial and terminal sigma of ranks,  $\sigma_{v'}$ , must be, as long as  $P$  is unchanged.

$$\sigma_{v'}^2 = \frac{P^2 - 1}{12} \quad (\text{Eq. 51c, Ch. X})$$

Eq. 51a is maximal when  $r = -1.0$  making

$$\max \sigma_{v''} = 2\sigma_{v'} \quad (\text{Eq. 51d, Ch. X})$$

Finding the percentage that Eq. 51b is of Eq. 51d gives Eq. 51a. Eq. 51a is but a special case of Eq. 47a, where the ranked data are, by definition of the units, in a rectangular distribution and cannot appear in the two all-or-none categories of a monopoly. Eq. 51a reaches 100% whenever the ranking of the competitors is completely reversed.  $C_p$ , being a sigma of a difference, is a function of the change in the dispersion and in the correlation (i.e., reranking) of the competitors. In the case of ranked data the dispersion ( $\sigma_v$ ) is held constant by the nature of the units, leaving only the reranking ( $r$ ) to determine the intensity of the competing,  $C_p$ .

A third special case is where the situation in some way limits the complete

transfer of the desideratum between competitors even in the maximal instance. Thus among pupils in school marked 100% for perfect work and 0% for none, no one pupil can accumulate the marks of all the others in a monopoly. Maximal effective competing is where all initial perfect marks become zero ones or vice versa. Here, instead of only two deviations of 100% being possible, P such deviations are possible, so that the 2 in the denominator of Eq. 47 must be replaced by P, resulting in  $C_p$  becoming identical with  $\sigma_v$  of Eq. 48. In this case the standard deviation of gains and losses in percentage marks is also itself the index of effective competing, expressing the percentage of maximal competing. For illustration:

<i>Pupil</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>etc.</i>
Marks on date $\tau'$	75%	70%	80%	80%	65%	etc.
Marks on date $\tau''$	74%	72%	80%	85%	65%	etc.
Gains or losses = $\tau V$	1	2	0	5	0	
Gains or losses squared = $\tau V^2$	1	4	0	25	0	$\sum \tau V^2 = 30$
	$(\sum \tau V^2 / P) \cdot 5 = C_p = 2.4\%$					(Eq. 52, Ch. X)

showing very small effective competing for higher school marks in this illustration.

34. A special case for all these secondary processes (competing, revaluating, and mobility), but particularly useful in the case of revaluating where the desires may be unmeasurable, is that case where the entities begin as equals and end up in a monopoly by one of them. Thus alternative desiderata,  $v$  in number, may be rated as equally desired at first, but on study one finally emerges as the best, and all the desire of the plurel becomes concentrated on it alone. Here the formula becomes a function of the number of desiderata  $v$ . The initial equal desire for each desideratum is the mean intensity of desire, namely,

$$\tau M_D = \frac{\tau D_{\Sigma v}}{v} = \frac{100\%}{v}$$

The standard deviation of the terminal "monopoly," recalling that the shifts balance each other (i.e., that  $\tau D_{\Sigma v} = 0$ ), is given by:

$$\sigma_D^2 = \frac{\tau D_{\Sigma v}^2}{v} \quad (\text{Eq. 54a, Ch. X})$$

Since the deviation of the one monopolistic (i.e., exclusive) desire is  $\tau D_{\Sigma v} - \tau M_D = 100\% - \frac{100\%}{v}$ , and of the  $v - 1$  other desires is  $\tau M_D$ , Eq. 54a, Ch. X becomes:

$$\sigma_D^2 = [(100 - 100/v)^2 + (v - 1)(100/v)^2]/v \quad (\text{Eq. 54b, Ch. X})$$

(Note that the small letter in the descript position denotes an aggregation and requires the summation sign before it to convert its meaning into a single number, a sum; but when standing alone, as in the denominator here, it means a single number without requiring the  $\Sigma$  sign.)

which on expanding simplifies to:

$$\sigma_D = M_D (v - 1)^{\cdot 5} = \frac{100}{v} (v - 1)^{\cdot 5} = \text{revaluating between values in the "equality-to-monopoly" case} \\ \text{(Eq. 54c, Ch. X)}$$

$$\sigma_D = 100v^{-\cdot 5} \text{ when } v \text{ is large so that } -1 \text{ is relatively negligible} \\ \text{(Eq. 54d, Ch. X)}$$

$$\sigma_V = M_V (P - 1)^{\cdot 5} = \frac{100}{P} (P - 1)^{\cdot 5} = \text{competing in the "equality-to-monopoly" case} \\ \text{(Eq. 55, Ch. X)}$$

and

$$\sigma_P = M_P (p - 1)^{\cdot 5} = \frac{100}{p} (p - 1)^{\cdot 5} = \text{mobility in the "equality-to-monopoly" case} \\ \text{(Eq. 56, Ch. X)}$$

The largest numerical value of this index (Eqs. 54, 55, and 56, Ch. X) is 50%, which is reached when  $v = 2$  (or  $P = 2$ , or  $p = 2$ ), for this initial "equality to terminal monopoly" is half the range from an initial monopoly to a terminal reversed monopoly. This formula is useful where the units are not observable but the critical conditions of initial equality and final monopoly can be observed. Thus, a woman buying a dress which she finally comes to feel is the *only* one that will do of the initially equally favored candidates, has revaluated to the extent:

50%	if among	2	candidates	( $v = 2$ in Eq. 54c)
40%	"	"	5	"
30%	"	"	10	"
14%	"	"	50	"

This illustrates the possibility of measuring exactly such intangible psychic phenomena where no units are determinate, as long as critical behavior points are determinable, such as initial avowed equality of feeling and final avowed exclusiveness, or concentration, of feeling.

These second-order processes and attendant relations of competition, mobility, and popularity have been defined in terms of standard deviations of shifts. They may also be defined in terms of the average deviation (A.D.) or other measures of dispersion of the shifts in the factors during a period. Substituting A.D. for  $\sigma$  in Eq. 48 and Eq. 49 gives Eq. 47b as:

$$_{AD}Cp = \Sigma(+_v V) = \text{index of effective competing in average deviation terms.} \\ \text{The sum is over either the } P \text{ persons or the } p \text{ plurals} \\ \text{(Eq. 47b, Ch. X)}$$

where  $+_v V$  denotes the positive shifts, i.e., gains in percentages. (Note that  $\Sigma(+_v V) = -\Sigma(-_v V) = .5\Sigma|+_v V| = .5(AD)P$ , since  $\Sigma_v V = 0$ .)

Similarly popularity may be expressed:

$$_{AD}Rl = \sum_{1}^{v,p} (+_v D) = \text{index of revaluating in terms of the average deviation. The summation is for either the } v \text{ desiderata to one plural, or for the } p \text{ plurals} \\ \text{(Eq. 53b, Ch. X)}$$

These indices in average deviation terms are simpler for persons not trained in statistics to comprehend, but as a basis for correlation and further statistical study the indices in standard deviation terms are usually preferable.

A note on terminology is in order. The secondary processes, which are defined by a formula of more than one base letter, have been assigned a two-letter symbol to denote the index measuring the process more conveniently:

Mobility or repopulating	= Mb	= $(.5\Sigma_t P^2)^{.5} = \sigma_{(\iota P)}(.5p)^{.5}$	(Eq. 57a, Ch. X)
Competing or regrading	= Cp	= $(.5\Sigma_t V^2)^{.5} = \sigma_{(\iota V)}(.5p)^{.5}$	(Eq. 57b, Ch. X)
Repopularizing or revaluating	= Rl	= $(.5\Sigma_t D^2)^{.5} = \sigma_{(\iota D)}(.5p)^{.5}$	(Eq. 57c, Ch. X)
Retensing	= Rt	= $(.5\Sigma_t E^2)^{.5} = \sigma_{(\iota E)}(.5p)^{.5}$	(Eq. 57d, Ch. X)
		(See S. 10 and 18, Ch. X)	
Reordinating	= Ro	= $\sigma_{(\iota I)} = (\iota I, .)^{.5}$	(Eq. 57e, Ch. X)

The form of the symbol for "reordinating" is more general than the four above it, as it is not expressed as a percentage of a maximum sigma as are the four processes above it. These four are all derived from the tension theory of societal action (Eq. 34, Ch. V and Eq. 35, Ch. X).

The term in the first column ("mobility—repopularizing") is more generally the conventional term for this kind of phenomena, while the term in the second column ("repopulating—reordinating") may be used to specify this kind of phenomena as measured by the specific indices, Mb, Cp, Rl, Rt, Ro, based on standard deviations.

In order that the student shall not forget that every process and every other term defined in this volume by a numbered equation is deducible as a special case of the basic S-theory equation (Eq. 7a, Ch. X), the deductive derivation of the index of competing may be sketched as an example. Starting from the tension theory matrix equation, as this has already been shown to be but a special case of the S-theory equation, the V is isolated on one side. Since a process is being considered, time has an exponent of minus one. In Brief-S notation the equation for a single period is:

$$\iota(PDE^{-1}) = \iota V \quad (\text{Eq. 47c, Ch. X})$$

The right-hand member denotes the gains and losses in period  $\iota$  of the desideratum that is being competed for. To summarize these gains and losses in a single index, their sigma is taken, derived by squaring both sides of Eq. 47c and computing the square root of the average square, giving:

$$\sigma_{(\iota PDE^{-1})} = \sigma_{(\iota V)} \quad (\text{Eq. 47d, Ch. X})$$

In order to reduce this sigma to more readily interpretable units as a percentage of the maximal possible competing, both sides are divided by the maximal sigma (as described above in deriving Cp) and, for simplicity, the symbol Cp

denotes the left-hand member. This gives the definition of competing, Eq. 47a above.

35.  $\sigma(I) = R_v =$  "revarying" of an index on many dates (Eq. 58a, Ch. X)  
(See S. 37, Ch. X)

36. A caution may be noted by the advanced student. Since this sigma is over a series of dates,  $R_v$  measures a process, and when divided by the overall time period it is one measure of the velocity of fluctuating, or instability. A change in the revarying is of the order of an acceleration, for here the difference in the velocity of revarying in two periods is found and this difference is divided by its appropriate time-interval, giving a time rate of change of the rate of fluctuating, and this is an acceleration.

Note that, as usual in mathematics, the operations within the parenthesis are to be performed first. Thus in redispersing  $\sigma(I)$ , the sigma on each of two dates is first calculated and then their difference, or change in the inter-date period  $v$ , is calculated. In reordinating  $\sigma(I)$  on the other hand, the changes in the index  $(I)$  are first calculated for each person (or plurel, or whatever may be the unit of frequency), and then the sigma of these changes is calculated. In revarying  $\sigma(I)$ , the aggregation of an index on  $v$  dates is first observed and then the sigma of these values is calculated. The general formula for redeviating (Eq. 41, Ch. X), covering all three,  $I$ , or  $I \cdot I$ , omits the parentheses and does not specify the sequence of operations.

37. The formula which includes both is:

$$v(\sigma I \cdot I) \quad (\text{Eq. 59, Ch. X})$$

which states a change in period  $v$  in the second moments,  $I \cdot I$ , between two indices, each expressed in sigma units. When  $I \equiv I$  (the main diagonal entries), Eq. 59 becomes the unit variances whose square roots are the standard deviations defining instances of redispersing by Eq. 45, Ch. X. When  $I \neq I$  (entries not in the main diagonal), Eq. 59 becomes the correlation coefficients defining instances of "recodispersing" by Eq. 60, Ch. X.

38.  $\pm v(\sigma I \cdot I) =$  "recodispersing," a change of a correlation coefficient  
(Eq. 60, Ch. X)  
(See S. 33, Ch. X)

39. This is the sister process with reordinating. Reordinating is measured by the sigma of changes; reco-ordinating by the correlation coefficient of changes in a population. Reordinating involves one changing index; reco-ordinating involves two changing indices. They are related again as the main diagonal cells to the non-diagonal cells in the matrix Eq. 43. The two together may be referred to as "ordinating," measured by a function of the second moment of changes. For "reordinating" this function of the second moment is the sigma, as usually this is more useful than its square, the variance. For reco-ordinating (a correlation coefficient) the covariance is converted to sigma units to put it into a  $-1$  to  $0$  to  $+1$  scale.

$(\sigma_t I)_{I \cdot J}$  = re-co-ordinating, the correlation of changes (Eq. 61a, Ch. X)

(See S. 34 and 47, Ch. X)

$(\iota I)_I^5 \cdot I, (\sigma_t I)_{I \cdot J} = \sigma(\iota I), {}^r(\iota IJ)$  = ordinating, the dispersion and the correlation of changes (Eq. 61b, Ch. X)

(Cf. Eq. 46a, Ch. X)

40. Also, see in this connection such a study as Ogburn's *Stationary and Changing Societies* (Ref. 51) for an excellent collection of hypotheses which can be quantitatively verified by appropriate choice among the indices measuring the redistributing processes, provided, as always, that data for computing those indices are securable. (The S-theory can be applied only insofar as data are available—which is a limitation common to all science, though the sciences vary in the availability of recorded data at present.)

41. This Brief-S formula is  $T^{-1} : I_i$  in full notation. For a system of culture complexes, subcomplexes, and culture traits a hierarchy of classification scripts, changing in a set of periods or dates,  $|_t$ , serves thus:

$I_{s/t} = (I)_{(1 : \dots : z)/t} =$  differential velocities of systems of culture changes (Eq. 62b, Ch. X)

Note that lead and lag is not limited to indicators. Other components may show differential velocities of change. Thus the length of the working day (a duration) may decrease less rapidly than the number of hours required to produce a given amount of goods. For an example of spatial lead and lag, the area of county government has lagged behind the area accessible in a day's travel.

42. This reliability is the significance ratio of a difference in means divided by the standard error of that difference:

$$\frac{M(I)_I - M(I)_J}{\sigma_{I-J}} = \frac{\Sigma(I)_I P^{-1} - \Sigma(I)_J P^{-1}}{(\sigma_I^2/P + \sigma_J^2/P - 2r_{IJ}\sigma_I\sigma_J/P)^{.5}} \quad \begin{array}{l} \text{significance ratio of a difference} \\ \text{between means of two indices} \end{array} \quad \text{(Eq. 63, Ch. X)}$$

For the reliability of an inter-plural difference, substitute  $|_{p'}$  and  $|_{p''}$  for  $|_I$  and  $|_J$ , and note that  $P_{p'}$  may not overlap with  $P_{p''}$ , so that the  $r$  term vanishes.

43. Just as reordinating is measured by a sigma of a difference (a time-change) and can be either directly observed or synthesized by Eq. 46b, Ch. X, so re-co-ordinating can be directly observed (by calculating the  $r$  of two series of observed changes) or synthesized by the formula for the correlation of sums and differences whenever all the necessary intercorrelation coefficients and sigmas are available.

44. The formula for reranking is usually:

$$v \cdot v'' I = \text{the correlation of an index on two dates} \quad \text{(Eq. 64a, Ch. X)}$$

But it may also include acceleration in the reranking of dynamic indices from two periods:

$$v \cdot v''(\sigma I) = \text{the correlation of an index from two periods} \quad \text{(Eq. 64b, Ch. X)}$$

45. For examples see Columns b, g, and k in S. 48, Ch. X. The size of the correlation measures the amount of constancy in the ranking; the amount of change in the ranking is given by its complement, the coefficient of alienation,

$(1 - \sigma \cdot \sigma' I^2)^{.5}$ , (see S. 12, Ch. VI). When the competitors in any field as in an athletic event are unchanged and their relative excellence ( $V$ ) is expressed in ranks, then the sigma cannot change on a second race, since  $\sigma$  in rank units is purely a function of  $P$  which is constant (see Eq. 51c, Ch. X, or Eq. 88 in S.17, Ch. VI). Then competing (or reordinating more generally) becomes a function solely of the reranking, as shown in Eq. 51.

46. Again, as in the case of revarying, if the indices are dynamic ones so that the series is one of periods rather than dates, then the recovarying process becomes of the order of an acceleration rather than a velocity. It is then a measure summarizing the tendency to change the velocities of the periods. Also a change in the recovarying, Eq. 65, of static indices is of the order of an acceleration.

Revarying measures the fluctuations of one index in time; recovarying measures the co fluctuating of two indices. They are related as the main diagonal cells to the other cells in the matrix Eq. 43.

47. The limits of these processes may be noted. They are 0 and the maximum of the index for redeviating processes, and either  $+1.00$  to  $-1.00$  for the recorrelating processes.

48.  $Sk$  = Pearson's  $\beta^{.5}$ . This  $Sk$  shows by its sign the sign of the skewing as  $\beta_1$  does not do. In symmetric distributions  $Sk = 0$ . As skewness may be  $+$  or  $-$ , the sign of skewing, its change in a period, must be carefully checked. Negative skewing,  $-_tSk$  is a change toward greater negative skewness or a smaller positive skewness; and positive skewing,  $+_tSk$ , is an increasing positive or a decreasing negative skewness.

Even though simpler measures of skewing based on differences between the median or mode, and the mean are used, for systematic purposes skewing may still be defined by the more accurate third-moment method and classed as a tertiary process.

49. They may not be completely operational definitions, if the indices entering into the formulae are not operationally defined. If the ingredients of a formula are merely descriptive concepts of undetermined reliability and validity, then the concept defined by that formula will be defined operationally in part only.

50. The four classes of spatial dimensioning may be symbolized in Brief-S formulae with suggested verbal labels as follows:

$L^0$ = change involving points = puncting	(Eq. 67a, Ch. X)
$L^{+1}$ = change involving lines = lining	(Eq. 67b, Ch. X)
$L^{+2}$ = change involving areas = "arealing"	(Eq. 67c, Ch. X)
$L^{+3}$ = change involving volumes = voluming	(Eq. 67d, Ch. X)

("Arealing" is a coined term, while the verbs "to punct" and "to volume" are rare forms found in Webster's Dictionary and revived here for technical denotation.)

51. The qualifying of space by societal qualities is represented by the attribute-space product (see Eq. 5, Ch. VIII), and is implicitly denoted by the class scripts on the spatial indices stating the kind of space denoted. Thus in  $L^2_1 (= I^0_1 L^2)$  of S. 61, Ch. X the class script  $|_1$  denotes "imperial," so that  $L^2_1$  denotes "imperial area," or "the territory of an empire."

# CONSPICUOUS OF SPATIAL PROCESSES (Note 52)

Spacing =  $\iota L^1$  (Eq. 67, Ch. X)

First Operation of Exponents, i.e., Spatial Dimensioning, $\iota(L')$		Puncting = $\iota L^0$ (Eq. 67a) Point Processes	Lineing = $\iota L^{+1}$ (Eq. 67b) Lane Processes	Arealing = $\iota L^{+2}$ (Eq. 67c) Area Processes	Voluming = $\iota L^{+3}$ (Eq. 67d) Volume Processes
Second Operation of Exponents, i.e., Spatial	( $L'$ )'' $T^{-1}$ , Binary processes (change of a two- factor index)	Spatializing = $\iota(L^1)^0$ (Eq. 67-0) Frequency- processes, i.e., zero- order moments Change in the num- ber of spaces	Linearizing = $\iota(L^{+1})^0$ (Eq. 67b0) Change in number of lines S. 9, Ch. II (if $T^{-1}$ ) S. 6, Ch. VIII (if $T^{-1}$ ) S. 11, Ch. VIII (if $T^{-1}$ ) S. 13, Ch. VIII $T^{-1}$ )	"Arealizing" = $\iota(L^{+2})^0$ (Eq. 67c0) Change in number of areas S. 7, Ch. VIII (if $T^{-1}$ )	"Volumizing" = $\iota(L^{+3})^0$ (Eq. 67d0) Change in number of volumes S. 15, Ch. VIII (if $T^{-1}$ )
		Spatializing = $\iota(L^1)^{+1}$ (Eq. 67-1) Mean-processes, i.e., first-order mo- ments Change in the size of spaces	Lineating = $\iota(L^{+1})^{+1} = \iota L$ (Eq. 67b1) Change in lineage S. 6, Ch. VIII (if $T^{-1}$ ) S. 60, Ch. X	Densating = $\iota(L^{+2})^{+1}$ (Eq. 67c1) Change in area S. 61, 62, 67. Ch. X S. 73, Ch. X	"Volumating" = $\iota(L^{+3})^{+1}$ (Eq. 67d1) Change in volume
		Respatializing = $\iota(L^1)^{+2}$ (Eq. 67-2) Variance-processes, i.e., second-order moments	Relining = $\iota(L^{+1})^2$ (Eq. 67b2) Redistributing of line-sects S. 9, Ch. II (if $T^{-1}$ )	Rearealing = $\iota(L^{+2})^2$ (Eq. 67c2) Redistributing of area	Revoluming = $\iota(L^{+3})^2$ (Eq. 67d2) Redistributing of volumes

Distrib- uting,	Change in the dis- tribution of spaces	$(L')''(I)^i T^{-1}$  Ternary * processes,  (change of a three- factor index)	$i(L')'' L^{-1}$ (Eq. 67L)    $i(L')'' P$ (Eq. 67P)    $i(L')'' I$ (Eq. 67I)	Change of points- per-space (Space = points, lines, areas, or vol- umes) (Eq. 67aL) S. 13, Ch. VIII (if $T^{-1}$ )  Change of person- points (Eq. 67aP) S. 11, Ch. VIII (if $T^{-1}$ )  Change of indicant- points (Eq. 67aI) S. 40, Ch. XI  Change of dura- tion-points (Eq. 67aT)	Change of lines-per- space (Eq. 67bL) S. 23, Ch. VIII (if $T^{-1}$ ) S. 59, Ch. X  Change of person- lines (Eq. 67bP) S. 59, Ch. X  Change of indicant- lines (Eq. 67bI) S. 59, Ch. X S. 68, Ch. X S. 38, Ch. XI  Change of dura- tion-lines (Eq. 67bT)	S. 6, Ch. VIII (if $T^{-1}$ )  Change of areas- per-space (Eq. 67cL) S. 44, Ch. XI  Change of person- areas (Eq. 67cP) Densing = $(PL^{-2})$ S. 63, 69, 70, 72, $T^{-1}$ Ch. X  Change of indicant- areas (Eq. 67cI) S. 69, 70, 71, Ch. X  Change of dura- tion-areas (Eq. 67cT)	Remapping = $(.5(c_p L^2)^2)^{1/3}$ S. 8, Ch. II S. 6, Ch. VIII (if $T^{-1}$ ) S. 66, Ch. X  Change of volumes per-space (Eq. 67dL)  Change of person- volumes (Eq. 67dP) S. 15, Ch. VIII (if $T^{-1}$ ) S. 73, Ch. X  Change of indicant- volumes (Eq. 67dI)  Change of dura- tion-volumes (Eq. 67dT)	S. 72, Ch. X

\* "Ternary" means that there are three factors in the index measuring the process, namely,  $I$ ,  $T$ , and either another  $L$ , or  $P$  or  $I$  or  $T$ . Every process is measured by a binary term — the entity changed and time being the two factors.  $I$  (A quaternary process is illustrated by S. 72, Ch. X with the quantic formula of  $T^{-1} I^{1/3} P^{-1}$ ). These ternary processes may involve any of the three orders of moments, though first-order moments (primary ternary processes) are most common and second-order moments (secondary ternary processes) are infrequent.

**52.** *Conspectus of Spatial Processes* (see pp. 726-727 for table).

For systematic purposes some new terms seem desirable (though usage will, of course, determine whether the terms proposed here are convenient or not). Most of the seemingly unfamiliar verbs such as "spatiating," "puncting," "voluming," "densating," "lineating," are rare forms found in Webster's International Dictionary (1915) and are here revived for technical purposes. The five terms in quotes are coined to have endings consistent with their row and column meanings—"arealing" = any change involving areas; "arealizing" = change of the number of areas; "volumizing" = change of the number of volumes; "volumating" = change of volume; and "densing" is taken from "densening." For consistency throughout S-theory the ending "-izing" denotes a nullary process ( $l^0 = 0$ ), "-ating" denotes (with some exceptions) a primary process ( $l^0 = 1$ ), and "re-" denotes a secondary process ( $l^0 = 2$ ). The secondary spatial processes may take all the redeviating and recorelating subforms as outlined for indicators, though to describe these at present seems hyper-refinement. Until it becomes evident that sociologists will find the major classes of this classification of spatial processes useful, the elaboration of its minor classes may be postponed.

The references to S-situations in the table are not all clear-cut examples of the process which they illustrate. In some cases these S-situations only record static data which, if reobserved later in time, would yield the process illustrated. These references are marked "if  $T^{-1}$ ." Some other illustrations are undeveloped and need to have sigmas or other indices calculated before they illustrate the process exactly.

**53.** Each of the ten combined dimensioning and distributing processes can occur:

- a. divided by  $L^1$  to express it as a ratio, or percentage, of some total of points, lines, areas, or volumes
- b. multiplied, or divided, by a population
- c. multiplied, or divided, by an indicant, or by time.

These combined processes, defined by ternary terms since they are indices with three factors ( $(L^1)$ ,  $(I)$ , and  $(T^{-1})$ ), are shown in the last four rows of the table above.

These four rows might be better represented as a third degree of the matrix in considering each row as a page duplicating the first four rows, except for the additional factor of that row. This yields 40 processes potentially (the 10 cells (of columns 1 to 4)  $\times$  4), of which 26 ( $10 + 16$ ) are indicated in the table above. The number of these 26 formulae, making up the subclasses of the spacing process, Eq. 67, permits specifying any cell, array, or degree of the matrix. Thus Eq. 67b means all lining processes (second column of the matrix); Eq. 67-1 means all primary processes ("spatiating" second row); Eq. 67P means all populational spacing (fifth row); Eq. 67c1 means "densating"; and Eq. 67cP can mean "densing," the change of mean persons per area (the cell in the third row and fifth column).

**54.** "Densing" is a simplifying of the dictionary term "densening." Increases may be called "condensing," and decreases "undensing," another term coined

to fill a gap. "Densening" connotes that the number of persons has increased in a constant space; but densing, being a new word, is free of this connotation and may mean change of persons, or of the space, or both, upwards or downwards, as long as their ratio changes.

Change of person-areas, Eq. 67c1 P/L, may alternatively be expressed either as a person  $\times$  area product (Eq. 67c1 PL), or as a change of the ratio of area per person (Eq. 67-1 L/P). This latter reciprocal of densing may be termed "sparsing," just as its static aspect, the inverse density described in Chapter VIII is "sparsity." For some purposes sparsing,  ${}_t(L^2P^{-1})$ , may be a more convenient index than densing,  ${}_t(PL^{-2})$ .

Note that the processes of I-densating (Eq. 67c1 I) may similarly be expressed as the ratio,  ${}_t(IL^{-2})$ , (Eq. 67c1 I/L), as its inverse,  ${}_t(I^{-1}L^2)$ , (Eq. 67c1 L.I), or as the product,  ${}_t(IL^2)$ , (Eq. 67c1 IL). The processes of T-densating may also be expressed similarly (except for the inverse  ${}_t(T^{-1}L^2)$  which denotes a change of velocity of change of area). Simple densating, being always an attribute-space product,  $I^0L^2$ , is equivalent to the inverse ratio  $L^2/I^0$ , since multiplication by an attribute is equivalent to division by an attribute, just as  $x^0y = 1y = y/x^0 = y/1 = y$ .

55. For examples suggesting some of these subforms see:

S. 10, Ch. VIII potential room for expansion of population by continents

S. 12, Ch. VIII densities of the various nations in 1920

S. 15, Ch. VIII a static situation of persons per room, which is a density and becomes densing as it changes in time

S. 63, Ch. X lynchings by States in 32 years

S. 67, Ch. X growth of territory by nations (area per specified plurel)

56.  $(.5({}_t\sqrt[2]{L^2})^2)_{\Sigma 1} \cdot 5 = Rm = \text{"remapping"} \quad (\text{Eq. 69, Ch. X})$   
(Cf. Eqs. 14a and 47, Ch. X)

57. The formulae for periodizing are:

${}_tT^{-1} : {}^1T^0 : {}^0T^0 = {}_tT^{-1} : {}^tT^0 = {}_t({}^tT^0) = {}_t\cdot {}^t = \text{periodizing, the list of dates (occurrences) in a series of periods}$   
(Eq. 70a, Ch. X)  
Full formula, I<sup>0</sup> explicit      Full formula, I<sup>0</sup> implicit      Brief-S formulae  
(See S. 38 and 52, Ch. X)

Its velocity is:  $\Sigma {}_tT = \text{velocity of periodizing, the number of dates per period,}$   
 ${}_t^0 = 91;0;0;0 \quad (\text{Eq. 70b, Ch. X})$

58. The two directions of during may be called "prolonging" and "shortening."

${}_tT^{+1} = \text{during} \quad (\text{Eq. 71a, Ch. X})$   
 $+{}_tT^{+1} = \text{prolonging} \quad (\text{Eq. 71b, Ch. X})$   
 $-{}_tT^{+1} = \text{shortening} \quad (\text{Eq. 71c, Ch. X})$

During may, of course, be for some quality, or person, or plurel, or for an aggregation of them, expressed as a mean during. It needs to be clearly distinguished from the populating process distributed by ages. The quantic for-

mula is the same for "durational populating" and for "populational during," but the difference is clearly shown in the descriptive formula where the dependent index,  $P$  or  $T^{+1}$ , whichever by the regular rules is written last, is the one changing. Thus a changing population pyramid has the formula  ${}^tT^{-1} : {}^tT^{+1} : P$  (Brief-S =  ${}^t{}_uP$ ), showing it to be a durational populating, while a changing expectation of life may be  ${}^tT^{-1} : \underline{P}_p : T^{+1}$  (Brief-S =  ${}^tT_p$ ), showing it to be a populational during.

This principle was induced from inspecting the situations of change involving both  $P$  and  $T^{+1}$  which the student may verify for himself.

<i>Populational</i>	<i>During</i>	<i>Durational</i>	<i>Populating</i>
(change of $T^{+1}$ classified by $P_p$ )		(change of $P$ classified by ${}^tT^{+1}$ )	
S. 2, Ch. II $\underline{P}_p : {}^tT^{-1} : T^{+1}$		S. 26, Ch. II ${}^tT^{-1} : ({}^tT^1 : P_p : a)_r$	
S. 80, Ch. X ${}^tT^{-1} : \underline{P}_p : a : T^{+1}$		S. 74, Ch. X ${}^tT^{+1} : ({}_pPT^{-1})$	
S. 85c, Ch. X $I_1^0 : M(T^{+1}P^{-1})$		S. 75, Ch. X $({}^tT^{-1} : {}^tT^{+1} : P_p)_{q,\Sigma p}$	
S. 10, Ch. XII ${}^tP : {}^tT^{+1} : \underline{T}^t$		S. 77, Ch. X ${}^tT^{-1} : {}^tT^{+1} : P_p$	
S. 12, Ch. XII $\underline{P}_p : {}^tT^{-1} : T^{+1}$		S. 79, Ch. X ${}^tT^{+1} : {}^tT^{-1} : {}_p\underline{P}_p$	
S. 15, Ch. XII $\underline{P}_{\Sigma p} : {}^tT^{-1} : {}_pT^{+1}$ (a per capita during)		S. 81, Ch. X ${}^tT^{+1} : {}^tT^{-1} : ({}_pPT^{-1})_p$ (a correlational durational populating)	
		S. 82, Ch. X ${}^tT^{+1} : ({}_p\underline{PT}^{-1})_p : a$	
		S. 42, Ch. XI ${}^tT^{-1} : a : {}_uT^{+1} : ({}_pPT^{-1})$ (an acceleration)	
		S. 43, Ch. XI ${}^tT^{-1} : {}_uT^{+1} : ({}_pPT^{-1})_p : a : r$ (an acceleration)	
		S. 11, Ch. XII ${}^tT^{+1} : ({}_pPT^{-1})_p$	
		S. 13, Ch. XII $({}^tT^{+1} : {}_pP, T^{-1})_p$	
		S. 14, Ch. XII ${}^tT^{-1} : ({}^tT^{+1} : {}_pP)_{p,\Sigma p}$	

In addition to these attribute-during and populational-during processes there are the less frequently recorded indicant-during and spatial-during processes. For an example of the latter, see S. 84, Ch. X, the change of travel time for a given distance from Pittsburgh to Philadelphia. If the inverse of travel time per distance were reported, it would be a linear velocity, a speed of travel, which, as it changed with the years, would become an acceleration.

The relation of during processes to acceleration is very close. The distinction depends upon the somewhat arbitrary rule given in Chapter IX for distinguishing a duration from a change, a collection of ages up to the present from a growth over a period. Durational processes have the quantic of  $T^{-1}T^{+1}$ , while acceleration has  $T^{-1}T^{-1}$ , i.e.,  $|^s = 91 j i l ; p$  for during, while  $|^s = 8 j i l ; p$  for accelerating. All these during formulae are ternaries of the form  $T^{-1}T^{+1}(I)$ .

For a situation, the formula of which is a quaternary, see S. 29, Ch. II with the quantic formula  $T^{-1}T^{+1}I^{+1}P^{+1}$ . No quinary ( $T^{-1}T^{+1}ILP$ ) during processes have yet been observed by us in the social science literature.

59. The detailed formulae for the secondary durational processes are found by substituting  $T$  for  $I$  in Eqs. 45, 46, 48, 60, 61, and 65, Ch. X, and adjusting scripts in the Brief-S formulae so as not to confuse durations and changes.

$\pm \sigma(\tau T)$  = durational redispersing (+ = dispersing; - = equalizing) the change in a period  $\tau$  of the sigma of a set of durations

(Eq. 72, Ch. X)

$\sigma(\tau : T)$  = rescheduling, the sigma of a set of changes in duration

(Eq. 73, Ch. X)

$\sigma(\tau : T)$  = durational revarying, the sigma of the durations on a set of consecutive dates

(Eq. 74, Ch. X)

The average deviation is simpler than the standard deviation and for some purposes may be preferable for defining these temporal processes.

$\sigma' \cdot \sigma'' : T$  = durational reranking or resequencing, the correlation between durations on two dates

(Eq. 75, Ch. X)

$\pm \sigma(\tau, \dots)$  = durational recodispersing, the change of correlation between 2 sets of durations

(Eq. 76, Ch. X)

$(\sigma : T), \dots$  = durational recovarying, the correlation between 2 kinds of durations in a series of periods

(Eq. 77, Ch. X)

In these formulae the "set" referred to is the total frequency of a distribution in which the units of frequency may be persons or plurels, or less often, indicators, spaces, or even durations.

For a single versatile situation capable of illustrating all the secondary temporal processes, study S. 83, Ch. X. The number of carfloats in a day is the amount of the periodizing process. The velocity of periodizing (i.e., the rate of occurring of carfloats per day) is 7 per day (except for a few days such as 5 on January 12). The duration of the interval between the loading of a carfloat by a locomotive and the towing away of it by a tugboat (time between solid and dashed lines of a pair in the graph) changes from one day to the next and is an example of during. The duration, which is a delay here, is either prolonged or shortened. These seven changes of duration may be averaged for the seven floats for any pair of days yielding a mean during. The mean during may alternatively be calculated by subtracting the mean duration of the seven floats on one day from the mean duration on the next day.

Next the sigma of these seven durations on one day may be subtracted from their sigma on the next day, thus measuring the redispersing process. The sigma of the durations of one float along the horizontal line for the 15 recorded days would measure the durational revarying, the constancy or fluctuations of the delays for any one float. Next, the sigma of the 7 differences between the corresponding durations in two columns measures the rescheduling, the variability of the schedule from one day to another. This could be averaged for all pairs of successive days to obtain a mean rescheduling. The correlation between the durations in two columns measures the reranking of the durations between two dates. The reranking and redispersing combine to give the rescheduling by the usual formula for the sigma of a difference. Recovarying is the correlation between a solid and its paired dashed line, the time of loading correlated with time of leaving in the time series of 15 days. Recodispersing can be illustrated, somewhat artificially, by shifting the duration from being the loading-to-leaving

interval to being the loading-to-midnight interval and the leaving-to-midnight interval. This yields two durations to correlate—the seven times of loading with the seven times of leaving on one day. The change in this correlation from one day to another measures the recodispersing. The effect of the blizzard of January 7 and of the Sunday lulls affects almost all the secondary processes. In summary in S. 83, Ch. X:

dispersion = sigma of durations in one column  
 redispersing = change of this sigma between 2 columns (i.e., dates)  
 rescheduling = sigma of changes between 2 columns  
 revarying = sigma of durations in one row  
 recovarying = correlation between 2 rows  
 reranking = correlation between 2 columns  
 recodispersing = change between 2 columns in the correlation between 2 durations in one column

60. This momentum is a “complex-single” process as it is measured by a single ternary index which always has at least three factors:

(I), some index denoting what changes

$T^{-1}$ , denoting the time of the change, an essential factor in the velocity ratio

P, the essential factor making a process become a momentum, which is defined as the product of a velocity of change and the number of people changed. It is the product of a speed of change and its extensity in a population.

61. An example of the western scientific hygiene culture complex, which had been reduced to measurement penetrating in this sequence into a backward district of Syria, is given in the author's *A Controlled Experiment on Rural Hygiene in Syria* (Ref. 12, p. 206).

62. The previous index of stability between generations measures the degree to which a characteristic is an inheritable one. Momentum here measures the extensity with which the characteristic is actually transmitted in one generation.

For a discussion of biological heredity vs. environment with the sterility of this issue as usually stated and a proposed restatement in terms of pre-conception and post-conception influences, see Lundberg's comments in Chapter XI of *The Foundations of Sociology*.

63. In using this momentum index, its *standard error* of sampling should be determined. When the population is constant, its formula is:

$$\sigma Mm = (\sigma I)PT^{-1} = \text{standard error of momentum (P unchanged)} \\ \text{(Eq. 87, Ch. X)}$$

where  $\sigma I$  is the standard error of the index of change determined by the usual formulae. (See Ref. 16.)

64. In such a summative combination the index of the process is a polynomial (binomial, trinomial, etc.), as contrasted with the single process which is defined by a monomial. In determining the relations of aggregated processes to any other processes, multiple correlation technics will be much depended on. In these cases the summative combination of the variables of the multiple team

in the regression equation indicates the utility of postulating a summative combination as a first approximation (assuming suitable units and origins). More accurate combinative functions should result from research.

65. Only the positive value of the square root of a variance is used in calculating Cp.

66. Of these two technics the second is probably more sound. If generally practiced, it will build a science based on measured entities whose relations are measured and hence can be systematized and predicted. The first technic involves the dubious assumption that these three processes are each wholly contained in exploiting, and that exploiting is an arithmetic sum made up of these three processes plus residual processes. This assumption is expressed in formula Eq. 43 (S. 7), Ch. VI, viewing correlation in terms of percentages of elements, common and specific. But it is not known whether this assumption is true in this situation, i.e., is  $j = 0$ ; or are there no specific components in any of the three processes which are not included in the sum "exploiting"? Eq. 42, Ch. VI is general and is a safe assumption, but here in proportion to the relative size of  $j$ , the specific elements, the multiple correlation coefficient can never reach unity. The "residual" processes then would have to be considered not as something to be added to the three, but, in part at least, as something to be subtracted from "coercing," "competing," or "dispersion" in order to reduce their specific elements to zero and make Eq. 43 of S. 7, Ch. VI hold true. This means that, in the search for residual components to achieve a deattenuated multiple correlation of unity, both negative residual components to cancel out specific elements in the predictive team of the three constituent processes and positive residual components to add to the team must be found, in order to measure the elements which were previously specific to the criterial or dependent variable "exploiting."

But these controversial and unproved assumptions are avoided by the second technic of calling "exploiting," *as measured by the rating scale*, one process, and calling the sum of the other three, as measured, another process, and proceeding to study the relations between these two processes, each of which is now defined and measured without resting on dubious assumptions.

The social scientist studies the strands of the complex web of our social life, seeking to analyze that he may then resynthesize the total pattern or any part of it. His first step is to fix on the "strands" whose interweaving he will try to trace. If he chooses a strand which proves to be not a singly acting unity, as he supposes, but a shifting composite, his findings will be confused. To measure a social entity is to define a "strand" unambiguously, preparatory to tracing out its relation to other strands and eventually to the whole pattern.

67. For a more refined mathematical formula for evolution in the field of biological phenomena and therefore expressed in indicators more observable in that field, see: Wright, Sewall, "A Statistical Theory of Evolution," *Jour. Amer. Stat. Assoc.*, Vol. XXV, N.S. 173A, March, 1931.

68. In addition to these eight examples of hypotheses towards reducing aggregate processes to measurable entities, a few others are sketchily listed below. The imaginative student, who is versed in S-theory and in the sociological literature, may readily extend the list. The invention of such hypotheses is relatively

easy compared to the vast amount of painstaking work, including organized collection of data, which is needed to verify the hypotheses and test their utility for purposes of prediction and control.

*Commercializing*—The competing of a minority for money is superposed upon the co-operating of another plurel, as when someone fences off and charges admission to some picnic spot previously enjoyed freely by the public.

$$+tV_{Co}; Cp ? = (I)_{Cm} \quad (\text{Eq. 98, Ch. X})$$

subtypes of "co-operating" and "competing" = "commercializing"

*Persecuting*—One party is coercively giving unequal treatment to another party supposedly its equal, as when one sect with political power legislates against another sect of fellow citizens.

$$t(^{\circ}V); Cr ? = (I)_{Pr} \quad (\text{Eq. 99, Ch. X})$$

"discriminating" and "coercing" = "persecuting"

*Stratifying*—When a population is dispersing into a hierarchy on some scale, and when the mobility between the levels approaches the vanishing point as in the castes of India of old, stratification exists. Stratifying is the hardening of such ordination. It is the increasing of superordination and subordination combined with making the barriers between levels more effective. The opposite trend, which might be termed "unstratifying" or "democratizing," is a combination of equalizing the status of the dispersed classes and increasing gross mobility. When Hindu outcastes are being newly admitted to temples and allowed to mix in other ways with the upper castes, "democratizing" describes the process.

$$+t(^{\circ}I); <Mb_G ? = + (I)_{Sr} \quad (\text{Eq. 100a, Ch. X})$$

"dispersing" and low "mobility" (i.e., less than average) = "stratifying"

$$-t(^{\circ}I); >Mb_G ? = - (I)_{Sr} \quad (\text{Eq. 100b, Ch. X})$$

"equalizing" and high "mobility" (i.e., greater than average) = "unstratifying"

*Sporting*—In sportsmanship, typified by competitive athletics, there is first of all a common desideratum promoted, such as "exercise with excitement,"  $+tV_{r}$ . As a means to this end competing is introduced. The sportsmen are striving to obtain all or most of the subordinate desideratum, "winning this game,"  $Cp$  for  $+V_{r}$ . Inevitably some lose, and the final test of sportsmanship is to be a good loser, to reduce one's desire for winning the particular game that is already lost,  $-tD_{r}$ . Poor sportsmanship results from lack of any one of the three constituent processes, as when the players cease to care most about co-operating in "getting on with the game," or cease to compete,  $Cp$ , either through losing zest towards winning, or through letting their zeal spill over into slugging and conflicting,  $-tP_v$ , or "grouch" about losing,  $+tD_{r}$ .

$$+tV_r; Cp_{r}; -tD_{r} ? = (I)_{Spr} \quad (\text{Eq. 101, Ch. X})$$

subtypes of "co-operating" and "competing" and "accommodating"  
= "sporting"

*Integrating*—Frequently several related processes are going on concurrently. Thus “integrating” might be a label clapped onto the four integrative processes where people are associating into fewer groups, are assimilating their characteristics, are accommodating their mutually incompatible desires, and are co-operating towards common desiderata.

$$-_t(|\Sigma_p); -_t(|\Sigma_i); (-_tD); (+_tV_{co}) ? = +_t(I)_{Intg} \quad (\text{Eq. 102a, Ch. X})$$

“associating” and “similarizing” and “accommodating” and “co-operating”  
= “integrating”

For further societal processes which have been proposed as concepts for Sociology, as well as for a review of classifications of societal processes, the student should study Eubank's chapter on Societary Action (Process) (Ref. 25, Ch. XIII).

## Chapter XI

### FORCES, $T^{-2}$

#### I. ACCELERATION AND DECELERATION, $(I)T^{-2}$ , BINARY CELERATION

##### A. Definitions

The last chapter explored the varieties of the *velocity* of societal change. The present chapter studies change of velocity, or more exactly, the time rate of change of velocity of a societal change. This is currently called acceleration, and is defined by a ratio in which time appears to the second power and appears in the denominator:

$T^{-2}(I)$  = an acceleration or deceleration of some index  
(Eq. 1a, Ch. XI)

$|^s = 8; i; l; p$  = the quantic number for any celeration  
(Eq. 1b, Ch. XI)

Societal phenomena in the class defined by  $T^{-2}$  belong in the quantic solid diagramed in S. 33, Ch. II, in the right-hand array of three piles of blocks from the right foremost to the rearmost pile. The discussion of this array will complete the systematic exposition of the quantic classification of quantified societal data which is provided by the S-theory. All of the quantic classification will then have been cumulatively discussed, sector by sector and exponential dimension by exponential dimension within each sector. The student will then be equipped to analyze any quantitatively recorded situation into its descriptive formula and classify it by its quantic number.

"Acceleration" has hitherto denoted positive or negative acceleration according to the increase or decrease of velocity. "Deceleration" is the term used by physicists for a decreasing velocity, but acceleration conventionally means either increasing velocity or changing velocity. To avoid confusion, it is here proposed to use the term "celeration" for the rate of change of

velocity, positive or negative, and reserve the term "acceleration" for positive celeration only.

A celeration is expressed as the number of units changing per period per period. Thus, an annual income is a velocity expressed as dollars-per-year, and its increase is an acceleration expressed as the increase of dollars-per-year per year. (See S. 26, Ch. XII.) The number of telephone calls per hour is the ordinate in S. 15, Ch. XI and measures the velocity of the process of telephonic communication in that plurel, while the rising and falling of the curve (the change of ordinate) measures the positive and negative celeration of communicating in units of telephone calls per hour per hour. Again in S. 2, Ch. XI the ordinate measures the velocity of the birth process in units of births per year, while the difference in any two adjacent ordinates measures the celeration of the process in units of birth per year-per year. Positive and negative societal celeration, acceleration and deceleration, are thus the speeding up and slowing down respectively of any societal process.

*B. Classification of Celeration by Temporal Pattern - Trends, Cycles, and Fluctuations*

Since celerations are a time series of dynamic indices, they may be classified, in common with all time series, as steady, periodic, or irregular, i.e., into trends, cycles, and fluctuations.

In the case of a trend, the time curve tends to rise steadily (or fall steadily) with the length of time within some specified period that is studied. The amount of the velocity (the ordinate as conventionally graphed) tends to vary with the duration of time (the abscissa). The velocity of the periods is correlated with the total duration up to each period. A formula for this is:

$$(I)T^{-1} \bullet T^{+1} \neq 0 = \text{trending celeration} \quad (\text{Eq. 2, Ch. XI})$$

This states that the scalar product of the velocity of change and the length of time is not equal to zero. The correlation coefficient is not zero; it is either greater than zero for an acceleration or less than zero for a deceleration. If the data are fitted better by some curve than by a straight line, the correlation coefficient is replaced by the correlation ratio for curvilinear correlation. Eq. 2, Ch. XI, states the amount of tendency towards a trend.

The slope of a straight-line trend is given by the regression equation derived from Eq. 2.

For an example of a steady trend see S. 10, Ch. XI, where the percentage circulation of religious periodicals in the United States has decelerated throughout the present century. For an example of a trend fitted by a curve note the excellent fit of the ogive of the normal probability curve in S. 5, Ch. XI, showing the initial acceleration and final deceleration in the enacting of compulsory school laws in the Northern and in the Southern States.

The accelerating trend of single culture traits is more easily observed than the trend of culture complexes, or of a total culture. These latter await the invention of complex indices for summarizing them in operationally specified ways, such as the various economic indices or the collection of societal indices in S. 23, Ch. XII, showing trends in the Soviet Union. Some single simple indices, however, may indicate major trends. Thus, the steady acceleration of patenting in the United States since 1870 indicates a trend towards a cumulatively complicated technological culture. The function of trends and cycles in prediction will be more fully discussed in the next chapter.

In the case of cycles there is correlation (usually curvilinear) between the societal velocity and the duration within each of the cyclic periods. This is the trend formula repeated for each of the cycles,  $\tau$ , namely:

$$\tau(IT^{-1} \bullet T^{+1} \neq 0), \text{ cyclic celeration} \quad (\text{Eq. 3, Ch. XI})$$

The period of one cycle is determined by finding that period which yields the highest correlations (in the parenthesis of Eq. 3).

For examples of seasonal cycles of celeration see S. 25, Ch. XII, where robberies rise in winter; S. 25, Ch. XI, where commercial failures wax and wane with the season; S. 19, Ch. XI, where malaria accelerates in summer and decelerates in the fall. For a daily cycle, note S. 15, Ch. XI, where telephone calls accelerate in the morning, drop at noon, rise again, and finally decelerate by evening. For business cycles of depression and prosperity see S. 17, Ch. XI, showing steel construction, and S. 25, Ch. XI, showing various economic indices from the depression of 1921 to that of 1933. For a life cycle study, see S. 7, Ch. XII, showing earning cycles of the poor. An institutional life cycle is dia-

gramed in S. 16, Ch. XII, and fuller discussion with hypotheses of such cultural cycles may be found in abundance in the sociological literature (see Refs. 6 and 66). The ascertaining of a cycle is the first step towards predicting that phenomenon's future course and modifying the cycle if desired (as in controlling malaria in S. 19, Ch. XI).

In the case of fluctuations, there is no correlation between the velocity of the societal change and the length of time it has been observed.

$IT^{-1} \bullet T^{+1} = 0$  = fluctuant celeration (without trend or cycles)  
(Eq. 4, Ch. XI)

Fluctuant celeration is also the residual celeration after the removal by statistical computation of any secular trend and cyclic celeration. If the fluctuations are major ones, i.e., short-time trends, they may be termed "lunge" and "lapse." What is a short-time and what is a long-time trend is relative, but it can be fixed by specifying the duration or the limiting dates in any given data. Thus, the income for foreign missions lunged ahead during the World War (see S. 14, Ch. XI) and lapsed slightly thereafter. These terms are not to be confused with cultural lead and lag which compare either purels or indices of culture traits, whereas lunge and lapse compare periods.

Whenever the acceleration of change is, (a) swift, (b) does not revert to its former condition, and (c) involves a large number of societal indices representing a large segment of the culture of a population, it is called a revolution. Thus, a political revolution, as in the United States in 1775-83, or the Industrial Revolution, are instances of sudden accelerations of changes which persist and which involve the whole political or economic culture. The general revolution of Soviet Russia since 1917, illustrates swift, apparently enduring, and a many-sided societal change. Whenever the acceleration of societal change and the number of characteristics changed are extreme deviants from the normal, it may be termed a revolution in the subsector defined by the indices of those characteristics:

$+^{>+\sigma}(T^{-2}I)_{i>50\%} = \text{a societal revolution}$  (Eq. 5, Ch. XI)

This formula specifies a societal revolution as a supernormal positive acceleration of change in a majority of characteristics of some

specified type. "Supernormal" was defined in Chapter V as greater than one sigma above the mean ( $> + \sigma$ ) in some frequency distribution. Here the distributions are of societal accelerations ( $T^{-2}I$ ) and of their extensities ( $|i|$ ). The fact that the revolutionary changes persist is the behavioristic proof of their being desired by the population undergoing the revolution, and justifies using the evaluative term "supernormal" rather than the objective non-evaluative term "majorate" as defined in Chapter V.

The function of a definition such as Eq. 5, Ch. XI, is to provoke better observing of societal phenomena. To determine whether any situation is a revolution, or what degree of deviation from normal it represents, requires the prior observing of many similar situations, reduced to indices, with distributions of their acceleration and their extensities. With such a background, an application of Eq. V to a given situation will yield more scientific "understanding" of it and a predictive probability for it.

### *C. Classification of Celerations by Sectors*

The index of the celeration may be homosectoral. The speeding up or slowing down of a process may be of an indicatory, a populational, a spatial, or a durational process. The graphed situations at the end of this chapter are grouped by sectors—first indicatory celerations, then populational ones, then their combination in distributional celerations (which are societal forces as will be seen below), then spatial and durational celerations, and finally celerations of such secondary processes as recorrelating and interacting.

In graphs of celerations the ordinates are usually velocities (i.e., dynamic indices,  $(I)T^{-1}$ ) and the abscissa conventionally represents time, so that a difference in any two adjacent ordinates measures the celeration in the period between those two ordinates. The slope of the curve is here the celeration. The slope is the tangent of an angle,  $(IT^{-1})/T$ , the ratio of ordinate to abscissa. But sometimes the ordinates are static indices, and with time as the abscissa, the slope measures the velocity. Here changes in the slope, alterations of curvature, represent celeration. In a smooth curve which is expressible in a calculative equation, the first differential of Calculus states the velocity at any given date

and the second differential states the celeration. Thus, the population curve of the United States, projected into the future by various methods of estimating in S. 19, Ch. XII, has static units as ordinates, but the changing slope of the curve shows acceleration first and then expected deceleration in the latter part of the present century. The Reed-Pearl logistic curve of population growth is a similar example. Pemberton's fitted ogive of the normal probability curve (S. 5, Ch. X), showing growth of membership in the National Congress of Parents and Teachers, measures by its slope the velocity of this adpopulating process, and by its change of slope the celeration of this process.

In the case of spacing (i.e., spatial processes), the rate of change of linear velocity is the physicists "acceleration." (See S. 39, Ch. XI.) But for societal purposes "acceleration" can as well be applied to changing rates of change either of areas (as in the annual crop acreage of S. 44, Ch. XI) or of volumes (as in daily per capita water consumption of S. 73, Ch. X).

In the case of timing (i.e., temporal processes) two subtypes are: (a) the rate of change of a changing duration as the travel time from Pittsburgh to Philadelphia (S. 84, Ch. X) since 1812; and (b) the rate of change of such processes as periodizing, re-varying, recovarying, and reranking, where the velocity of the process is somewhat obscured in the index defining the process. The rule is simply to observe the time rate of change of the indices defining these processes.<sup>1</sup> \*

## II. SOCIETAL FORCE, $T^{-2}(I)P$ , TERNARY CELERATIONS

### A. Definition of "Force" and Examples

The necessary foundation of concepts has now been laid so that the superstructure defining a *societal force* can next be constructed. This concept is much used in the sociological literature. The journal of *Social Forces* evidences this use in its title. In other social sciences the concept is used with appropriate adjectives, such as an economic force, a political force, an educational force, etc. A force is conceived as something that does something—and that is about as definite as the notion often is. It is here

\* For Eqs. 6a-d, Ch. XI, see notes at end of chapter.

proposed to define again,<sup>2</sup> operationally, and with mathematical exactness the concept of a force in the social sciences.

A societal force is here defined as the celeration of change in a population. A force is measured as the product of the celeration and the population celerated:

$$F = T^{-2}(I)P = \text{a societal force} \quad (\text{Eq. 7, Ch. XI})$$

For an example of a societal force consider the economic force of taxation. The annual tax rate in dollars per year in the case of a poll tax is the velocity,  $IT^{-1}$ , of this particular taxing process. If this rate is changed, its annual rate of change of velocity is an acceleration,  $IT^{-2}$ . If, finally, it is multiplied by the number of persons taxed, one has a measure of the total force,  $T^{-2}IP$ , of this tax in the specified population and period. The force varies with the number of taxpayers and with the annual tax rate and its rate of change. Increases in this are positive forces; decreases in this are negative forces. The force is here expressed in units of person-dollars per year per year.

For an educational force consider a school requiring five course units per year of its students ( $IT^{-1}$ ). This is one rough measure of the velocity of the educational process as defined in this situation. If the school raised the requirement in one year to six courses, the process would be accelerated one course per year per year ( $T^{-2}I$ ) in that particular year. If this is multiplied by the number of students thus accelerated one has one measure of the educational force in that situation, expressed in units of student-courses per year per year. To teach more students means a greater educational force; but also to teach a larger body of knowledge in a given time means a greater educational force (or alternatively, to teach a given body of knowledge in less time means a greater force).

For a force of public opinion, consider an attitude test (I) which yields a score reflecting the intensity of a person's feeling on some public issue, such as an immediate declaration of war on another nation. Suppose the war fever is rising and with the test repeated at monthly intervals the velocity of increasing war enthusiasm is measured; an overt act by the enemy which suddenly increases this velocity in one month, produces an acceleration. If instead of an average score the total score of all persons

taking the attitude test is used, this sum of variable scores of  $P$  persons is equivalent to a product of  $P$  and a constant score (i.e., the average score). The acceleration of the total score, expressed in score units per month per month, is here one measure of the force of public opinion pushing that nation towards the war. Again the force varies with each of its three constituent factors: (a) the number of people stirred up, (b) the intensity of their feeling, and (c) the speed, or shortness of the time.

This definition of a force in Eq. 7, Ch. XI is for an *effective* force, since it is defined by its observed and measured effect. It is also a net force, since it is that which accomplishes the observed changing over and above whatever resistances there may have been in the situation. This definition is not an analogy from Physics, although it is based on analogous reasoning.<sup>3\*</sup> The reasoning is based on the principle that science as systematized and verified observations can be best verified by different persons, if its systematizing concepts are *operationally* secured from objective observations. The index is the record of objective observations.<sup>4</sup>

The formula, Eq. 7, specifies the operation of twice dividing the index by the time and multiplying by the population. By providing an objective and operational definition of a societal force, Eq. 7 should make the concept of a force more determinate and, therefore, more useful for prediction and control of dynamic societal phenomena.

### *B. A Societal Paraphrase of Newton's Laws of Motion*

Before proceeding to analyze societal forces by exploring the units in which they are measured or, to synthesize forces into systems, the assumptions of this definition of a force should be exposed. These assumptions may be conveniently expressed, for societal phenomena, by using historical precedent and paraphrasing Newton's three "laws of motion" for physical phenomena. This paraphrase, it should be clearly understood, does not assert that physical laws carry over into societal realms. The paraphrase is viewed by the author as not stating a law of nature or of society at all, but simply as a generalization which proposes an opera-

\* For Eqs. 8a-b, Ch. XI, see notes at end of chapter.

tional definition of the concept "force," in the form of a frame of reference by which a force may be quantitatively determined.

The paraphrase is:

1. Whatever changes the status of a population, or its process (changing), in rate or direction is called a societal force.

2. The societal force is proportional to the rate of change of the rate of societal change, and takes place along the line defined by the index measuring the change.

3. Forces and their total resistances are equal and opposite.

In terms of S-theory, where a static index measures the status of a population and a dynamic index,  $(IT^{-1})$  measures the rate of its changing, and where every index defines a direction (a vector) in societal space, the first two propositions can be more compactly stated as:

1. A societal force is whatever changes an index of some population.

2. The force is proportional to the celeration of the index and acts along the vector defined by the index.<sup>5</sup>

The first proposition simply proposes a name for the speeding up or slowing down of change in some respect in a population, and proposes the existing condition as a zero point from which to start measuring amounts (or directions). Instead of the *existing* state of static or changing phenomena any other state could be taken, but then the problem would be to choose which of innumerable possible states to take as the standard zero point. The state existing when our observing starts is unique, it is operationally defined, and it is objective in that it secures agreement between different observers as to what to call the zero point. There is no metaphysical implication that the present condition is the normal nature of society; it is merely an arbitrary but highly useful convention to standardize zero points in the flux of societal phenomena that seethes around us in a myriad of shifting patterns, so that we may standardize our measuring of societal forces relative to such zero points. It is comparable to adopting the freezing point of water as the zero point on the centigrade scale for temperature. Such a convention serves well until, for any given characteristic, further research supersedes it by the discovery of a more natural, or more logically compelling, or more constant, or even an absolute, zero point for that particular characteristic.

The second proposition proposes units for measuring amounts and directions of a force from the zero point. It proposes that the unit of force be proportional to the celeration of a unit of the index ( $IT^{-2}$ ). The most convenient proportion is unity, which makes the unit of the index the basic factor in the unit of force. The other factors are time and the population which is only implicitly suggested in such phrases, in the propositions above, as "change . . . of a population." Accordingly these propositions are less exact than the definition of a force given by Eq. 7, Ch. XI. The unit of force in its various forms is further discussed in the next section of this chapter.

The last part of the second proposition defines a force as a vector quantity, i.e., as having both amount and direction. The direction of a line is measured by units of angle, such as the 360 angular degrees of a circle, from another line taken as a standard of reference. Whenever such a standard reference vector is lacking, as it is still lacking in most sets of societal data, the next best alternative is to measure directions relative to the other vectors by plotting the angles between every pair of indices in a specified situation. These angles are specified (as described in Chapter VI) by their cosines, which are operationally secured by calculating the correlation coefficients between the two indices. This angle between the two vectors representing two forces measures the degree of qualitative difference-to-similarity of the two forces. It measures the extent to which they overlap and the residual extent of their independence. It measures the extent of their common component and the extent of their different components. The vector then specifies the qualitative kind and the quantitative amount of a societal force.

The third proposition makes it possible to write an equation balancing all the positive and negative forces of a given kind (i.e., along a specified vector) that are acting in a situation. Often some of these component forces are not identifiable, much less measurable with our present technics, and then the equation is only an hypothesis by the aid of which we can reason about and sometimes manipulate the forces, known and unknown, in the total situation. For some simple examples of the balancing of explicitly recorded societal forces consider S. 38, Ch. XI, where the rise or fall of monthly railroad revenues, expenditures, and

net revenues are three forces related in a system where the third is the difference of the first two:

$$s/M_0^2p'(F, - F_{,,} = F_{,,,}) = \text{the equation of business forces} \\ (\text{Eq. 9, Ch. XI})$$

This states that the income forces, less the resisting expenditure forces, equal the profit-or-loss forces. The forces are expressed in the same units of dollars per month per month for one business plurel. This equation is the essence of all business forces. The essence of business, as far as it is measured in monetary units, is to maximize the income forces, to minimize the resisting expenditure forces, in order to maximize the positive difference as a profit.

For another typical example of explicitly recorded balanced forces consider S. 8, Ch. XI, which shows the positive force of apartment building in the United States exactly compensated by the negative force of house building. The exact compensation is due to expressing each in units of percentages of their sum, as thus any decrease of one part of the whole is necessarily compensated for by the increase of other parts. (For other examples of this percentages-of-the-whole type of forces see S. 9, 11, 12, 20, 21, 31, and 37, Ch. XI.)

### *C. Units of Force*

#### 1. THE PROBLEM OF COMPARABLE UNITS FOR DISSIMILAR FORCES

The problem of the units for expressing a societal force requires study. The unit of societal force is defined as one index-population-unit per period per period,  $PIT^{-2}$ . It is a compound unit derived by multiplying the index units times the population units and dividing twice by the time units. It is specified by specifying the units of its three factors, just as the dyne, the unit of physical force, is defined as one gram of mass, times one centimeter of length, twice divided by one second of time.

The time factor in the societal force has usually one year for its unit, but multiples of decades and centuries and submultiples of months, weeks, days, hours, minutes, etc., are also frequent. Occasionally more irregular units, as of one era, one dynasty, one generation, one weekend, one inning, etc., may occur and yield a unit of force which is correspondingly vaguely delimited.

The units of the index which measures that characteristic of the population or its environment which is changing, are of many kinds. There are millions, perhaps, of such characteristics, as noted in Chapter III, when classifying indicators by content. The great diversity of indicators and compounds in indices presents a problem here for it means that societal forces will be expressed in many different units, making comparison difficult. This is a sharp contrast to physical forces where no such problem occurs.

Towards solving this problem of comparable units for societal forces the following classes of comparable index units have been induced from our collection of situations (cf. Ch. III, Section IV B):

Attribute units,  $I^0$ , "P-forces"

All-or-none units, percentages

Ordinal units, ranks

Cardinal units, such as:

Money

Standard deviations

Index numbers

Components and epsilon elements

Each of these will be discussed next.

## 2. ATTRIBUTE UNITS, $I^0$ , "P-FORCES"

When the index is at the primitive limit of being an attribute,  $I^0$ , a simple qualitative characteristic, its product with the population,  $P$ , in Eq. 7 merely characterizes that population qualitatively. Thus, if "immigration" is the attribute  $I^0$ , then immigrants  $P$ , ( $= I^0P$  by Eq. 4, Ch. IV) is its product with a population. The unit is "one person immigrating." This attribute-population product includes all peopling processes,  $\{P^p$ , so that the celeration of any peopling process is the simplest type of societal force. Since the attribute is usually implicit the formula for a force appears to reduce to celeration of a population.<sup>6\*</sup> It is the change of the number of persons, or plurels, of some kind. Hence this type of force may be called a "peopling force," or simply a "P-force." It is of common occurrence as in population studies, where changes of birth, death, marriage, divorce,

\* For Eq. 10, Ch. XI, see notes at end of chapter.

migration, and other dynamic population rates are instances of P-forces. (See S. 4, 5, 8, 42, 43, Ch. XI.)

Wherever persons are the unit of population they provide comparable units for the P-forces built on them; but where plurels are the population unit, the units of different P-forces will be comparable only insofar as the plurels are comparable.

### 3. ALL-OR-NONE UNITS, $^1\text{I}$ , PERCENTAGES, $\%$ I

The next type of comparable unit is the all-or-none attribute, a quality observed as either present or absent. By the convention of using the easily understood percentage units, each case of the presence of the attribute is assigned the numerical value of 100, and each case of its absence is assigned the numerical value of 0. Every percentage ever calculated is a mean of such an all-or-none variable as explained in Chapter III. By using percentages many diverse sorts of characteristics can be expressed in comparable units, and the units of force based on them become comparable. (See for examples, S. 9, 10, 11, 21, 31, 32, 33, and 37, Ch. XI.)

### 4. ORDINAL UNITS, RANKS, $^1\text{I}$

At the next stage of precision in observing phenomena, rank units may be used. These vary from crude comparisons of "more" or "less," to formal ranking in assigning first, second, third places, etc., on to percentile units which are ranks transposed to a scale ranging from 0 to 100. Rating scales and all manner of daily subjective judgments and preferences, made by everyone, utilize such ordinal units. Much of this use is not conscious, and may not even be verbalized to the extent of such a thought crossing one's mind as saying, "I want to do this now more than that," and thereby assign "this" a rank of 1 and "that" a rank of 2. But such implicit judgments are easily converted into formal ranks and provide comparable, though not very accurate, units for very diverse sorts of phenomena. Societal forces expressed in ordinal indices seem rare, however, in the sociological literature, as no such situations have turned up in the collection of 1500 situations. They are easily conceived, however, as in ranking competitors on each of three dates. Changes of rank between two consecutive dates give the velocities, and the change of the two velocities yields an acceleration which, multiplied by the popula-

tion, is a measure of this societal force of competition in rank units. Forces of competing in athletics can then be compared with forces of competing in scholarship, or in some other characteristic.

#### 5. CARDINAL UNITS, I, MONEY, INDEX NUMBERS, SIGMAS, COMPONENTS, AND EPSILONS

More accurate units than any of the preceding are cardinal units which are standardized, equal, and interchangeable. Five subclasses of this cardinal class may be noted as partial solutions to the problem of comparable F-units (as units of societal force may be conveniently labeled for short).

The first subclass is that of monetary units in all its forms. To exchange and compare the great diversity of goods and services wanted by man, barter proved inadequate and a common unit of money has been developed. Thus, in the equation of business forces above (Eq. 9, Ch. XI), and in S. 38, Ch. XI, the common unit of dollars enables comparison of all the kinds of things which brought in income and all the kinds of things which required expenditure, and enables getting their difference. In vectorial terms every item of income, or expense, is a vector whose direction is determined by all the characteristics of that item, but only the projection of that vector on a common money vector is considered in this situation. The algebraic sum of these positive and negative projections is the profit vector. This vectorial interpretation of projections onto a common vector holds for all types of common units and is not limited to monetary units. (See S. 14, 16, 20, 38, Ch. XI.)

Index numbers are ratios, simple or compounded in various ways, to compare amounts changing in a time series, as well as to compare different kinds of things. By dividing the amount of some characteristic in each period by its amount in some one period which is taken as the base and assigned a value of 100, the amounts of change and derived velocities, momenta, accelerations, and forces of this characteristic become comparable with those of other characteristics which have been similarly reduced to an index number. (See S. 13, 25, 26, 32, 44, Ch. XI.)

A third type of cardinal unit which makes diverse data comparable is the standard deviation (and other similar units of dis-

persion, such as the average deviation, probable error, coefficient of variation, etc.). Measurements of human characteristics from weight to intelligence, from income to musical talent, from cephalic index to religious attitudes, from basal metabolism to school achievement, and a host of others, can be compared when recast into "standard scores." Correlation coefficients are expressed in standard deviation units. For celeration of change of human characteristics standard deviation units are perhaps the most important technic for increasing comparability. (See S. 34, 35, Ch. XI for examples.)

A fourth technic for comparability is the recently developing field of component analysis as described in Chapter VI on Correlation. Indices may be analyzed into their principal axes, centroids, or other types of components. Forces based on such indices are correspondingly analyzed thereby into component forces of those types. (These techniques may utilize sigma units, if the matrix of intercorrelation coefficients is taken as the starting point of the computation of components, but the matrix of second moments may also be the starting point and thus not involve sigma units.)

A further extension of intercorrelational component analysis is the determination of the epsilon elements, described in our hypothesis of multiplex elements in Chapter V and further developed as the  $\epsilon$ -elements hypothesis in Chapter VI. If with further research these common elements of correlated indices were to become more precisely determinable, they should provide comparison par excellence. It is conceivable that the utility of the atom as a unit in Chemistry may become rivaled by the utility of such epsilon elements for the analysis and resynthesis of indices of societal characteristics in common units.

## 6. COMPARABLE P-UNITS, "I-FORCES"

Paragraphs 2-5 above classify the ways for comparing societal forces as regards the index factor in the formula for a force,  $T^{-2}(I)P$ . The units of the time factor have been noted. Consider next the units of the P factor.

The population factor in the societal force usually has one person as its unit. It may also be one plurel, but in this case the plurels must be comparable if forces are to be compared. Thus,

forces expressed in national plurals may be compared with other forces built up out of national plurals, but not of family plurals, nor of "talkie" audiences.

A very frequent case is that of a single plural. If the plural is the unit and there is only one of them, the formula of the force practically reduces to a celeration of the index. The term "I-force" is suggested to name this class of forces.

$$F_i = T^{-2}(I)P^0 = T^{-2}(I), \text{ an "I-force"} \quad (\text{Eq. 11, Ch. XI})^7$$

This is the counterpart to the P-force of Eq. 10, Ch. XI. In an I-force the P factor is unity leaving I as a variable, while in a P-force the I factor is unity ( $I^0 = 1$ ) leaving P as a variable. An I-force states a celerated index in some one plural; if there is an aggregation of plurals, there is an aggregation of I-forces. (For I-forces in one plural see S. 14, 17, 26, 31, 33, 38, 44, Ch. XI; for aggregations of them, see S. 9, 15, 18, 19, 20, 21, 22, 32, 37, Ch. XI.) For implicit I-forces which are written as celerations only (since the plural is not even named in the situation as recorded), but which actually refer to some plural, see S. 9, 11, 12, 13, 24, 25, 27, 39, 40, Ch. XI. Most celerated situations recorded in the literature of the various social sciences are either of the P-force or of the I-force type; the full IP-force has seldom been computed.

Another frequent case closely related to the single plural of an I-force is the case of a force for a single person. If only one person's celeration has been observed, as in S. 23, Ch. XI, the production curve of an employee when in training, the situation is a psychological one. But when a plural has been observed, the plural number makes the situation a sociological one, even though subsequent division by the population reduces the force to a per capita force for one person. But this is the average person typifying the plural. Such a force in per capita units may be called a "per capita force," and symbolized by  ${}^pF$  denoting the force for each person of an aggregation.

$${}^pF = T^{-2}(I_P)P^{-1} = T^{-2}\sum_1^P(I)P^{-1} = T^{-2}(MI) \text{ a per capita force} \quad (\text{Eq. 12, Ch. XI})$$

The second member of this equation asserts the celeration of an index for a plural in per capita units; the third member asserts

explicitly that the index for the plurel is a sum of the indices of the persons in that plurel; the fourth member asserts that the celeration is of a mean index. These express the same kind of thing but differ in their explicitness. In one sense such a per capita force by holding the P factor constant at unity tends to make forces more readily compared, as then only the I factor has to be transposed to comparable units. But in another sense the per capita force is more a celeration than a force. The essence of a societal force is its being a product of an I and a P; if then it is divided by P so that the P cancels out, a celeration of an index ( $T^{-1}I$ ) remains, which is not a full societal IP-force. Hence, this force should be distinguished by the adjective "per capita," denoted by the person script,  $\mathfrak{p}$ , in Eq. 12 above. Per capita forces are illustrated in S. 25, 28, 29, 30, and 35, Ch. XI.

One further point on F-units will guard the student from confusion. It should be noted that the product IP which assumes a constant (i.e., a single valued) P, multiplied by a single valued I, is equivalent to the sum of the variable indices of the P persons. Every linear product is a sum in the special case where each factor has but one value for a given value of the other. Geometrically, the sum is the area of any frequency distribution, the product is the area of a rectangular frequency distribution, i.e., a distribution where there is a constant frequency of the distributed units (P) at every class-interval of the distributing variable (I). This equivalence of sum and product is at once evident if the mean index (the fourth member of Eq. 12, Ch. XI) is multiplied by the population, giving the sum of the indices in the population (the third member of Eq. 12 with  $P^{-1}$  canceled out). All this signifies that an index for a plurel can be regarded as a sum of the per capita shares of the phenomenon measured by the index, and hence as equivalent to the product of an index and a population of persons. Thus, two hospitals for a town of 10,000 people can be considered either as the sum of the per capita shares of hospitalisation (namely .0002 hospitals per person), written as

$\sum_{10,000} (.0002) = 2$ , or as the product of the shares times the population ( $.0002 \times 10,000 = 2$ ). In this sense an I-force (Eq. 11, Ch. XI) can often be regarded as implicitly a full IP-force. The I-force has one plurel as its population unit; this can be resolved,

implicitly at least, into a full IP-force with persons as the population units. For concrete examples of this principle, study S. 15, an aggregation of two similar I-forces of telephonic communicating, in each of two district plurels, where the total number of calls is equal to the product of the population of the district and their calls per capita ( $P^M I = \sum_1^P \sum_1^P I P^{-1} = \sum_1^P I$ ), and S. 18, Ch. XI, where the total number of patents, giving an I-force, in each nation is equivalent to the product of the population and the per capita patenting.

In summary, the unit of societal force is a compound unit defined as one population unit changed one index unit per period per period. The two time periods need not be the same (as in S. 40, Ch. XI, where they are decades and years, or S. 39, Ch. XI, where the time units are years and hours respectively). More concisely a unit of societal force is one party (= a person or a plurel = "P-unit") times one index unit ("(I)-unit") per period per period. It is the celeration of one PI unit,  $F = T^{-2}IP$ . This unit may conveniently be named an "F-unit," meaning a unit of societal force. Since it is a societal dyne, the coined term "sodyne" has been suggested. In earlier publications the author suggested the term "stim," meaning one unit of stimulation (Ref. 12, p. 212, and Ref. 18, p. 60). As societal forces are not identical with the physiologists' and psychologists' concept of a stimulus, however, this may not be the best term. As it has not, to the author's knowledge, been used by any other sociologist<sup>8</sup> since the suggestion was published over five years ago, the more objective name, F-unit, is now recommended in keeping with the notation of S-theory as being more flexible for the addition, in the future, of units of any other operationally defined concepts symbolized by letters. Thus, the unit of momentum can be called one "Mm-unit" ( $Mm = T^{-1}IP$ ), and one acceleration unit can be called one "A-unit" ( $A = T^{-2}(I)$ ). (Physics has no name for either its unit of momentum or of acceleration.)

#### *D. Systems of Forces*

In any quantitatively recorded societal situation, S, there may be more than one force acting. Whenever the data for calculating these forces are explicitly available the problem of combining

these forces arises. The technics of combining them have been fully developed in vectorial algebra. Forces may be added, subtracted, multiplied, and, if collinear, they may be divided into each other. To do this, each force is represented by a vector. The length of the vector is the number of F-units, and the direction of the vector represents the qualitative nature of any force relative to other forces. To add forces, their vectors are added by placing them end to end and drawing a line from the beginning of the first vector to the tip of the last vector as explained with diagrams in Chapter III. To subtract forces, the negative vector of the force to be subtracted is added. Physicists call adding forces the "composition of forces" into a "resultant" force which is their sum, and the subtraction, or analysis, of forces, the "resolution" of a force into its "components." These components can be identified with the components of intercorrelated indices, discussed in Chapter VI, whenever the indices are in F-units. These technics of intercorrelational component analysis show how observed forces may be resolved into independent components. In general, the angle between any two force vectors is given by the cosine, which is calculated as the correlation coefficient between the two forces. If, as is usual, the two forces to be combined have the same time units and population units these are constants, and the correlation of the forces is the same as the correlation of their indices of change (I). This correlation coefficient is the scalar product of unit vectors and is, therefore, the technic for multiplying two forces together, as diagramed in S. 13, Ch. VI.

A force whose vector is collinear with another is either a "facilitating" force if similarly directed, or a "resisting" force if oppositely directed. Collinear forces are perfectly correlated together, either positively or negatively. A force whose vector is perpendicular to a second force-vector is "neutral." Any other force whose vector makes some angle between  $0^\circ$  and  $90^\circ$  with a second force-vector is a "redirecting" force. If added to the first force its direction will be changed. Forces which operate on each other, i.e., which facilitate, resist, or redirect each other, will necessarily be correlated. If forces are uncorrelated, it is evident that they probably do not operate on each other. This is only probable and not certain, for there may be some influence on each other, some correlation, which is obscured or neutralized

by other forces. The manipulation of a force to achieve socially desired changes has the double aspect of reaccelerating and re-directing it, of changing it in quantity and in quality. Thus, in S. 38, Ch. XI, typifying the financial forces in any business, the income and expenditure forces are collinear but oppositely directed. Expenditure is a "resisting" force to the income force. Profits are the net force remaining after subtracting the resisting force from the income force in any pair of months of this time series. In S. 34, Ch. XI, where prices and church increases are negatively correlated to the extent  $r = -.67$ , the vector of an I-momentum (price changes in the U.S. plurel) and the vector of an I-force (church acceleration in the U.S. plurel) make an angle of  $132^\circ$  with each other. This should not be superficially interpreted as that prices cause church increases (or vice versa). Each index reflects, and perhaps quite indirectly reflects, a combination of interworking economic and psychological forces whose net effect is merely indicated by the indices of price and church increases. The accurate statement is that a momentum is correlated with a force—the momentum, the correlation, and the force are accurately defined, whereas causal relationships must remain as further inferences from the facts until more complete data as to causality are presented.

Before exploring the relation of "causes" to "forces," however, the technic of determining the completeness of a system of forces should be noted. Suppose that any one force,  $F_0$ , taken as a criterion for analysis, is analyzed into its component forces  $F_A \cdots F_N$ . On adding the forces, is their resultant,  $F_{\Sigma n}$ , the criterial force? Statistically this operation is calculating the multiple correlation between the criterial force and the set of  $n$  forces.

$$\begin{aligned} F_{\Sigma n} &= F_0 && \text{(Eq. 13a, Ch. XI)} \\ \text{if } F_{\Sigma n} \bullet F_0 &= 1 && \text{(Eq. 13b, Ch. XI)} \end{aligned}$$

If the correlation is perfect, these components are all the necessary and sufficient ones. By as much as the correlation is less than unity there are unknown, or unmeasured components of  $F_0$ , which are not included in the set  $F_n$ . The extent of the unknown components is measurable as a percentage by the coefficient of non-determination,  $k^2$ , which is the square of the coefficient of alienation ( $k$ ):

$$100k_0^2 \cdot A \dots N = 100(1-r_0^2 \cdot A \dots N) = \% \text{ coefficient of non-determination}$$

(Eq. 14, Ch. XI)

(See S. 12, Ch. VI)

This is the lack of correlation in percentage terms; it is a measure of the degree of unmeasured residual correlation between the criterial force and the set of  $n$  other forces. This technic has been extended by Hotelling to cover the case of "double multiple" correlation between two sets of variables. (Ref. 30.) The double multiple coefficient of non-determination from this technic measures the incompleteness with which either set of variables expresses the other set. When the variables are forces, this coefficient states the degree of incompleteness with which the forces of either set can be considered to be component forces of the other set of forces.

This coefficient of non-determination may often serve as a crucial test of the Gestalt theory, the theory of configurations, in the social and psychological sciences. The more dogmatic proponents of the modern Gestalt theory, which originated among German psychologists of this century, assert that a "whole" living phenomenon is more than the sum of its parts. The pattern, the form, the gestalt, is something over and above the additive combination of the parts and is the essence of a living "whole." The definition of a "whole" is not always clear, nor is the distinction between the permitted resolving of a whole into smaller wholes and the non-permitted resolving of a whole into its parts. The theory emphasizes the importance of the arrangement, of the configuration, of figure and background, or situational setting, in studying phenomena. In somewhat oversimplified terms, the crucial test of the theory is whether a specified whole phenomenon is or is not the sum of its parts, as 6 is the sum of 2 and 4 without remainder. This test can, wherever appropriate data is securable, be quantitatively carried out by calculating the coefficient of non-determination between the specified whole and its alleged parts, whether these are forces or other kinds of correlatable variables. If the coefficient  $k^2$  is zero, that "whole" is exactly the sum of its parts without remainder, for then these parts will duplicate perfectly, or will predict with

100% certainty, that "whole." In proportion to the size of  $k^2$  that specified whole has an unmeasured remainder over and above the sum of those alleged parts. Since in most measured societal phenomena,  $k^2$  is greater than zero (i.e., correlation is seldom perfect), the Gestalt theory would seem to have much evidence in its support when applied to the current very inadequate data and analyses of situations.

Whenever a  $k^2 \neq 0$  (within standard error limits) is found, it is a challenge to further research to overcome that measured amount of our ignorance. It should never be taken as a metaphysical mystery that there will always be an unknown residuum, for that is an unscientific dogma. There are many cases where science has so adequately factored a phenomenon that their sum yields the whole phenomenon perfectly. The causation of certain diseases by specific germs and their elimination by inoculations and drugs, the generation of a whole baby by specific causative-prior acts and conditions, and many other phenomena have been so adequately factored into operationally defined component entities that their recombination will, whenever desired, yield the whole specified phenomenon. The key to the Gestalt issue is the *adequate* factoring of a whole into parts. An arch is a whole which is inadequately factored into 11 stones, as the sum of 11 stones may not make an arch. But when factored into 11 stones whose shapes and relative locations are specified by a blueprint, the arch can be perfectly synthesized as the sum of *these* factors. The gestalt, the form, must be included among the factors if the factoring is to be adequate. Just so in societal phenomena the Gestalt theory should call attention to the current inadequate factoring of phenomena, rather than be considered as a metaphysical ultimate principle of reality. The scientist should measure the inadequacy of his factoring of, for instance, a societal force into component forces by calculating  $k^2$ —and then set to work to reduce  $k^2$  towards zero by more adequate factoring, by researching for data which more completely describe, in operational, and therefore duplicable terms, the force at issue.

#### E. "Causes" vs. "Forces"

In dealing with societal forces much confusion will result unless "a force" is sharply distinguished from "a cause." The concept

of causation has many connotations and much indefiniteness, to eliminate which, the concept of forces is useful. A societal force is the measured acceleration of change of a population. A societal cause is an antecedent condition which, under stated conditions, has a certain probability of accelerating change of a population. A cause involves the three elements, as defined in Chapter VI, of a time sequence, a probability (in either its simple form, or some compounded form such as correlation—see Chapter VI) and the specification of the conditions. The force is an effect; it is defined by the observed change which has been effected. Its cause must be antecedent to it in time, and must either be correlated to it or have a probability between 0 and 1 of being followed by it, under specified conditions. In the social sciences it is often difficult to isolate and correlate antecedent conditions in the total on-going flux of societal phenomena, so that thinking in terms of causes is often apt to be in terms of inferences and subjective guesses, while thinking in terms of forces, the measured accelerations of people's changing, is thinking more in terms of objective facts. There is a danger, however, that the sociologist may tend to use the concept of "forces" as synonymous with "causes." The force is the accelerating change, the cause is that which precedes it with a certain probability.<sup>9</sup>\*

In this connection the term "agent" is convenient for denoting the people and/or material things which are thought to be the cause, or part of the cause of a given effect. The doctor and the hospital which cure a patient are agents, whereas the cause may be an act, such as the operation of removing a gangrenous appendix, or conditions facilitating internal processes, such as rest in a tubercular sanitarium. The agent, like an engine generating a physical force, is the persons and material things whose functioning may be the cause of accelerating some societal change. Thus, the army and its equipment are agents, human and material, whose functioning in war is the direct cause of the conflicting forces which are measurable by the effects in casualty units, i.e., by the acceleration of the depopulating process. (See Eq. 10, Ch. X.)

The term "factor" is sometimes substituted for a "force" or a "cause" in the current social science literature. In S-theory a

\* For Eq. 15, Ch. XI, see notes at end of chapter.

factor has a strictly mathematical meaning as any part which is *multiplied* by other parts to make a whole. A factor may be a force, a population, an attribute, an area, or any index, if it is multiplied by some other index. A "component" is allied to a "factor" in being an index representing any part of a specified whole where the best mathematical method of combining parts may be unknown, but where addition is taken for a first approximation as in adding the weighted variables of a multiple regression equation. Since the components, whether forces or not, are measured by indices which are vector quantities, the addition is vectorial addition. The addition adds qualities and quantities by adding directions and lengths. Factors, then, are entities combined by multiplication, components are vectorial quantities combined for a first approximation by addition. Another property of components is that when multiplied to form the scalar product, which is the correlation coefficient, they are seen as correlates. Every correlate is a factor in a scalar product. A cause is a correlate and hence a factor as defined here. But not all factors are causes. Causes are a special subclass of factors.

In discussing forces and their causes, the relations of these concepts to the concepts of societal "tension" and "desire," defined by our value theory, Eq. 35, Ch. X, should be exposed. Tensing and desiring are processes which, if accelerated in some population, become forces. They may be causes of societal phenomena, if they precede those phenomena with a certain probability under specified conditions. The usual sequence is for the desire to celerate first and, depending on the celeration of the desideratum, for the tensing to celerate correspondingly. This tensing force stimulates people to activity tending to accelerate progress, the increasing of the desideratum. Thus, in competitive phenomena the index of competing was defined as effective competing, the final redistribution of the desideratum. Causative competing would seem to be correlated to the antecedent desiring and tensing which motivated actions resulting in the redistribution of the desideratum, V.

#### *F. The Standard Error of a Societal Force*

In calculating a societal force its standard error must be computed. Forces cannot be legitimately compared until it is deter-

mined whether an observed difference in their number of F-units is statistically significant, or whether it may be due to random errors of sampling. The standard error formulae for societal force as well as for its constituent concepts of velocity, acceleration, and momentum have been derived for both the simple case of a constant population and for the complicated case of a population which varies from survey date to survey date. For derivation of these formulae, discussion, and numerical illustration of special cases see Refs. 16 and 18, and Appendix III.<sup>10</sup>

### III. SOCIETAL CONTROL, $T^{-2}IP^2$

A sociological concept closely allied to that of societal forces is "societal control." This concept is much used in the sociological literature; it is proposed here to give it a more exact and operational definition. In S-theory, control is defined as the combination of the interacting of people with a societal force. More exactly, "societal control" is the acceleration of change of one plurel by another plurel. It is specified by the formula:

$$S_{cn} = {}^pP_p :: {}^pP_p : {}_tT^{-1} : IT^{-1} \quad (\text{Eq. 33a, Ch. XI})$$

$$|s = 8;1;0;2 \quad (\text{Eq. 33b, Ch. XI})$$

This is not an index (I) as is a force, F; it is an aggregative situation (S), since the  $P^2$  symbolizes an interrelational matrix of numbers and not a single number. It means that parties (= persons or plurels) are cross-classified with some accelerating index of a changing characteristic in each cell expressing the interacting of each pair of parties.

An explicitly recorded situation fulfilling this definition of societal control is S. 37, Ch. XI, which records the celeration of annual trade ( $T^{-2}I$ ) between one national plurel and each of five other plurels ( $\underline{P}, :: \underline{P}_q$ ). The accelerations and decelerations of trade between any two years, multiplied by one plurel, is an I-force (see Eq. 11, Ch. XI). Trade is a dynamic characteristic of a stimulus-response type between parties, and therefore, falls within the definition of "interacting" in Chapter VII. This sub-type of societal control is the speeding up or slowing down of an economic interacting between parties.

This definition of societal control is objectively observable wherever a response of parties which is directly related to the

stimulation of other specified parties is observable. The operation of observing and recording the characteristic, the response, in some index, in dividing it by the time period to give the velocity, in dividing it a second time by the interval periods to get the acceleration, in multiplying it by a population to get the force, and in cross-classifying it in a matrix with other parties to establish the interaction—all these operations together constitute the operational definition of "societal control." The merit of such a definition is that different observers can apply the formula and reach agreement as to whether a given situation, or set of statistics, involves societal control or not. The current definitions of societal control in terms of descriptive synonyms, etc. are less objective and agreement-compelling. This last sentence is an hypothesis. It can be put to the test of a crucial experiment; it need not be a matter of opinion and argument; it can be proved, or disproved, to be a fact. The crucial experiment would be to select several hundred recorded situations which a majority of a panel of sociologists rate as: (a) involving societal control in x situations, (b) involving societal force in y situations, (c) involving both in z situations, and (d) involving neither in w situations. With this body of data to test the hypothesis, let a different panel of sociologists, equally well trained in S-theory and in any other school or textbook's theory of "control," classify each situation as to whether "force," "control," both, or neither, are involved, first on the basis of one theory, then on the basis of the other. Whichever theory yields the highest percentage (or other index) of agreement between the members of the second panel is thereby proved to be the most objective.

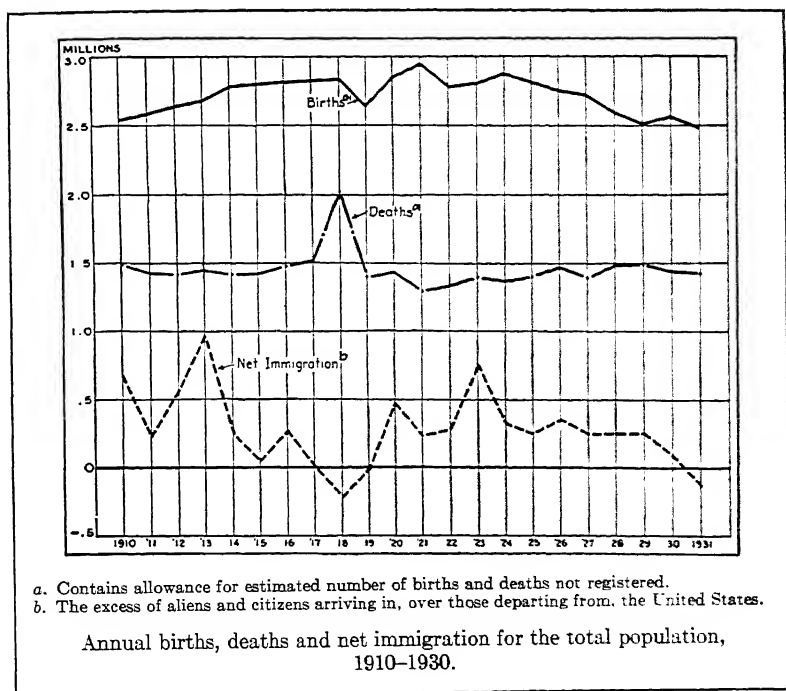
This tests the reliability of determination of alternative concepts; it does not test their utility or immediate significance for science. The working hypothesis, however, is that in the long run science will build more firmly and become more powerful in dealing with phenomena in proportion to its being based on reliable concepts, even, if necessary, at a sacrifice of immediate apparent significance. Unreliable concepts rise and wane with emotional fashions; reliably determinable concepts become the more enduring foundation stones of a science. As long as "a force" in Physics remained the anthropomorphic concept of a wish or a will of some person or spirit, it had little scientific utility. When it was

operationally defined as a measured acceleration of some mass, physical science laid a permanent foundation stone on which further building could be based. As long as "societal control" has a mystic or at least intangible element in Sociology it remains of limited utility; when it is operationally defined as a measured celerating of interacting of parties, Sociology may build on it.

This definition of societal control neither necessarily implies, nor yet excludes, the notion of desirability of the control. The value judgment, the social approval or disapproval, the desire or aversion of people of specified plurels for the celerated interacting is a separate issue. This desirability of the interacting can be factually observed by speech or other behavioral indicators of people's desires, and the question of whether the control represents progress or regress to a specified plurel at a specified time can be determined. But the concept of control is evaluatively neutral.

Control is any celerating interaction, whether good or bad, consciously wanted or not. If a crime wave grows, criminals are increasingly controlling society; if police reduce it, they are controlling the criminals. If villagers who eagerly desire health but are ignorant of any germ theory teach their children to drink filthy ditch water as healthier than well water, and a water-borne epidemic breaks out among those children, those villagers are controlling their youth towards disease, although they neither know it nor want it. Consciousness of effects and ethical evaluation of effects are humanly important, but science in the long run will build more surely and give man greater mastery of these effects if it uses concepts which, while they may bear on ethical issues, are objectively definable with less of subjective indeterminateness.<sup>11</sup>

## S. 1



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill Book Company, Vol. I, 1933, p. 38.

Descriptive formula:  $S_1 = {}^tT^{-1} : (PT^{-1})_p$

Quantic number = 8;0;0;1

Legend:

$S_1$  = The situation  
 records for each of

$|_p = 3$  plurals  $\left\{ \begin{array}{l} \text{births} \\ \text{deaths} \\ \text{migrants} \end{array} \right.$

${}^tT^{-1} = 21$  years

$P$  = the number of persons

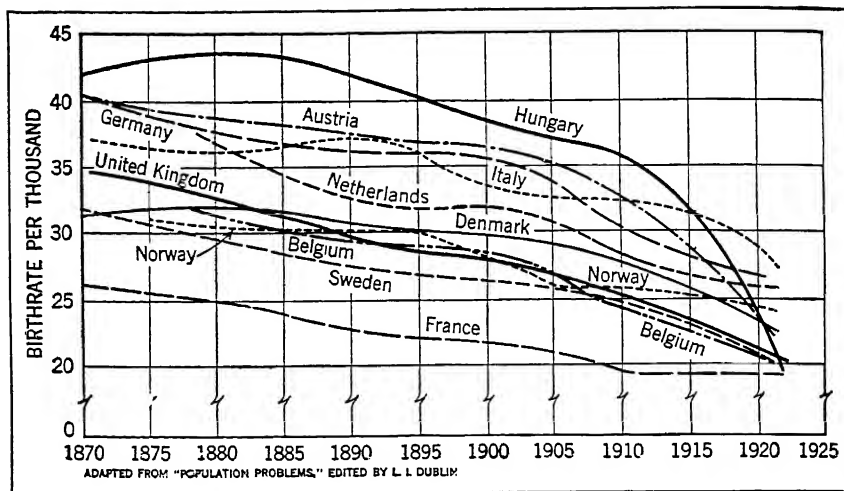
' = beginning in 1910  
 for each of

$T^{-1}$  = per year

Comment:

The graph gives specific data for the United States for the general equation analyzing population change into its four components (Eq. 11, Ch. X). The curves also show the fluctuations in the vital forces measured in units of persons, per year per year, born, dying, or migrating. Writing these three implicit attributes explicitly gives the formula for these three peopling forces ( $P$ -force,  $T^{-2}T^0P$ )<sub>i</sub>.

## S. 2



Ref.: Ross, Edward Alsworth, *Principles of Sociology*, The Century Company, 1930, p. 25.

Descriptive formula:  $S_2 = \{T^{-1} : (\%PT^{-1})_p$

Quantic number = 8;0;0;1

Legend:

$S_2$  = The situation

records for each of

$\{T^{-1}$  = 55 years

'| = beginning in 1870

$T^{-1}$  = the annual

$\%P$  = birth rate

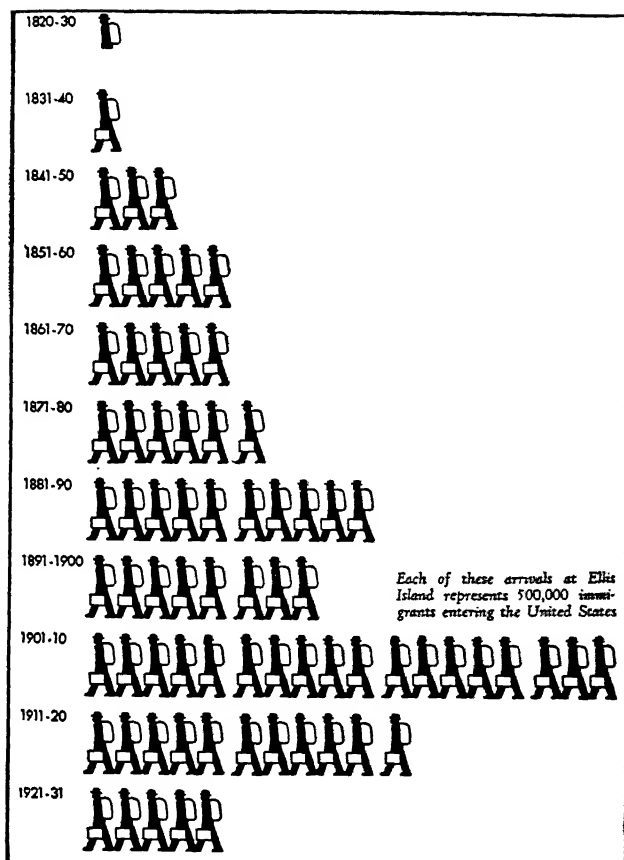
for each of

$|_p$  = 14 European nations

Comment:

The downward slopes of the curves measure the decelerating of one form of the adpopulating process. The societal force is here a P-force whose unit is .001 person born per year per year. The percentage sign indicates relative units which are here tenths of one percent.

## S. 3



Ref.: Neurath, Otto, "World Planning in U. S. A.," *Survey*, Vol. LXVII, No. 11, March 1, 1932, p. 622.

Descriptive formula:  $S_3 = {}^tT^{-1} : (PT^{-1})$   
 Legend:

Quantic number = 8;0;0;1

$S_3$  = The situation  
 records for each of

'| = beginning in 1820  
 P = the number of immigrants

${}^tT^{-1}$  = 11 decades

$T^{-1}$  = per decade

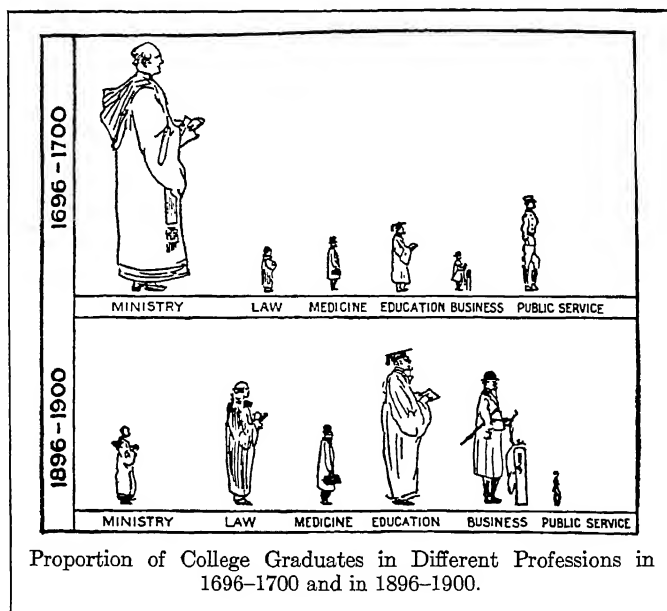
Comment:

This is a "P-force,"  $T^{-2}I^0P$ , where the all-or-none qualitative characteristic  $I^0$ , is whether a person was an immigrant or not. Strictly the attribute, implicit in P as usual, is an all-or-none one. It has the values 0 and 1 for non-immigrants

## S. 3 (Continued)

and immigrants respectively. Thus non-immigrants vanish leaving only the number of immigrants,  $I^0P_0^{+1} = P$ , where  $I^0$  = the immigrating attribute,  $P_0^{+1}$  = pure or unqualified persons, and  $P$  = immigrants.

## S. 4



Ref.: Brinton, Willard C. *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 38.

Descriptive formula:  $S_4 = ,T^{-1} : (\underline{P}T^{-1})_p$

Quantic number = 8;0;0;1

Legend:

$S_4$  = The situation

in

records

$T^{-1}$  = a 4-year period

$,T^{-1}$  = the change in 2 centuries

who enter each

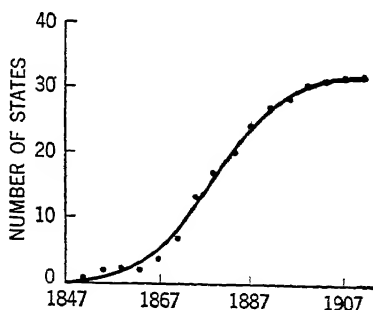
in

of

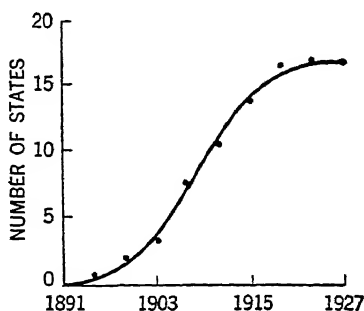
$\underline{P}$  = the number (which is not stated) of college graduates

$|_p = 6$  professions

## S. 5



Dates of enactment of compulsory school laws by northern and western states; four year intervals, 1847-1911. Fitted with the cumulative curve of a normal frequency distribution of which the  $\sigma$  is 11.76 and  $Y_0$  is 4.2.



Dates of enactment of compulsory school laws by southern states; four year intervals, 1891-1927. Fitted with the cumulative curve of a normal frequency distribution of which the  $\sigma$  is 6.72 and  $Y_0$  is 4.05.

Ref.: Pemberton, H. Earl, "The Curve of Culture Diffusion," *American Sociological Review*, Vol. I, No. 4, Aug. 1936, p. 553.

Descriptive formula:  $S_s = ({}_tT^{-1} : ({}_pPT^{-1}))_p$

Quantic number = 8;0;0;1

Legend:

$S_s$  = The situation  
records for each of  
 $P_p$  = 2 plurels in the U.S.A.  
during

${}_tT^{-1}$  = 20 4-year periods  
 ${}^1$  = from 1847 onwards  
 ${}_pP$  = a frequency of States adopt-  
ing compulsory school laws  
 $T^{-1}$  = per quadrennium

Comment:

This situation illustrates a P-force, i.e., a societal force where the indicator is an attribute so that the characteristic is all-or-none. The population unit

S. 5 (*Continued*)

here is a plurel, one State. The unit of force is one State adopting these laws per year per year. Expressed in normal probability terms, the general calculative formula for both situations is:

$$S = (t : \Sigma p| = \frac{\Sigma p|}{\sigma \sqrt{2\pi}} e^{-T^2/2\sigma^2})$$

where  $t : \Sigma p|$  is the frequency of adopting States in each quadrennium,  $\Sigma p|$  is all the States in that situation,  $S$ ,  $\sigma$  is the standard deviation of the distribution, and  $T$  is the deviation from the mean date ( $e = 2.7$  the log base). This reduces to the following two equations for normal probability distributions (from which the ogives above were cumulated):

for the Northern States:

and

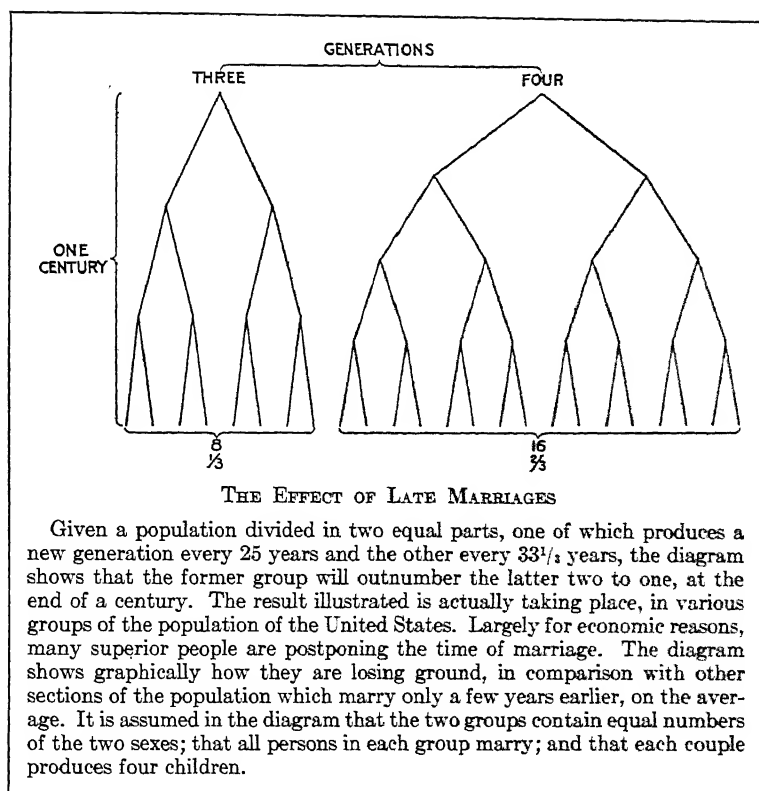
for the Southern States:

$$t : \Sigma p| = 4.2/2.7^{T^2/27}$$

$$t : \Sigma p| = 4.05/2.7^{T^2/90.31}$$

This situation illustrates the possibility of discovering regularities in cultural phenomena, such as to be describable by rational equations. These equations are not purely empirical ones, such as in a lucky finding that a normal curve fits the data well. The normal curve was selected for fitting because it is created by the same sort of elements which were assumed (by the hypothesis to be tested) to have created these law enactments—namely, a large number of relatively small and independent causes, or elements of influence.

## S. 6



Ref.: Popenoe and Johnson, *Applied Eugenics*, Macmillan, 1933, p. 108.

Descriptive formula:  $S_8 = (v; uT^{-1} : PT^{-1})_p$

Quantic number = 8;0;0;1

Legend:

$S_8$  = The situation  
records in

$P$  = the increase of population  
 $T^{-1}$  = per generation

$vT^{-1}$  = one century

for each of

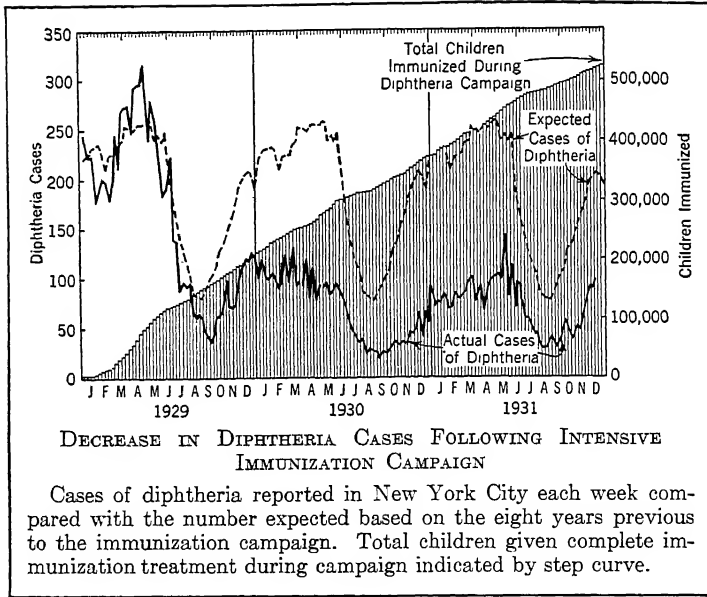
$u$  = subdivided into generations

$|_p = 2$  reproduction plurels

Comment:

The societal force is here a P-force as usual in situations describing the accelerating growth of a population. The units are one person (P) generated ( $I^0$ ) per generation ( $T^{-1}$ ) per generation ( $T^{-1}$ ).  $F = T^{-2}I^0P$ . The attribute which qualitatively characterizes the population is implicit as usual in the attribute-population product  $I^0P_0 = P_1$ .

## S. 7



Ref.: Boldman, Charles F., "Downing of Diphtheria," *Survey*, Vol. LXVII, No. 8, Jan. 15, 1932, p. 420.

Descriptive formula:  $S_7 = (\tau : u' T^{-1} : PT^{-1})_p$

Quantic number = 8;0;0;1

Legend:

$S_7$  = The situation

$T^{-1}$  = per year

records for each of

in each of

$\tau T^{-1}$  = 3 years

' = since 1929

$: u$  = subdivided into months

P = the diphtheria cases

$|_p$  = 3 plurels

{ actual,  
expected,  
immunized  
cumulative

Comment:

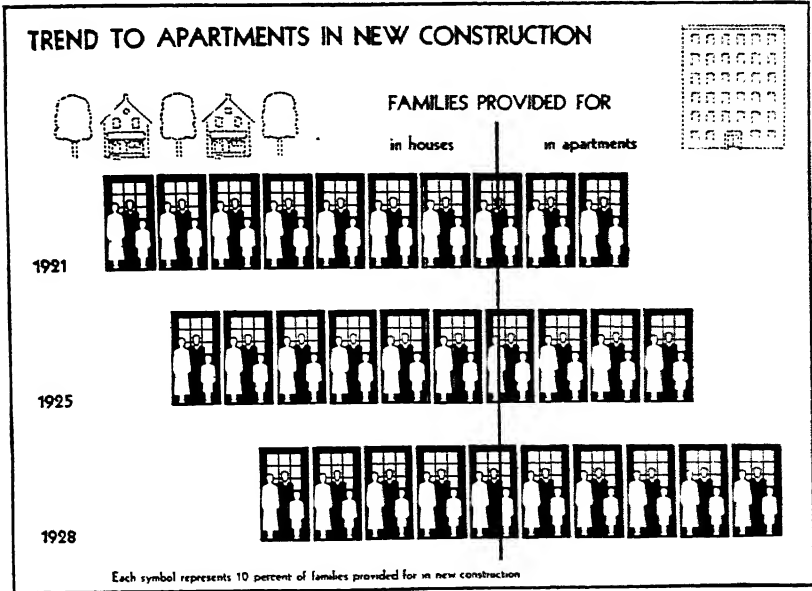
The situation provides the data for the analysis of a societal force into its component forces. The fluctuations of the curve of "actual cases" shows the effective P-force which is the vectorial resultant of all facilitating and resisting forces. The "expected" curve shows the forces facilitating diphtheria (of which seasonal influences are a major subcomponent), and the "immunized" curve shows the force resisting diphtheria and reducing the "expected" to the "actual" curve.

This is an implicit illustration of our definition of societal control which is the accelerating of change in one plurel by another plurel. If the Department of Health and co-operating medical profession were explicitly recorded in the situation, they would be the controlling plurel to be cross-classified with the

## S. 7 (Continued)

children who are the controlled plurel in the typical descriptive formula for explicit societal control,  $T^{-1} : P, :: P,, : I^1 T^{-1}, (1^1 = 8;1;0;2)$ . In the situation as recorded here  $|^1 = 0$  and  $|,, = 0$  (i.e., not mentioned, a "nul" class), so that this typical equation reduces to  $T^{-1} : P, T^{-1}$ , which is the formula given above (without the additional qualifying descripts,  $: u_p$ ).

## S. 8



Ref.: *Research Bulletin of the National Education Association*, Vol. XII, No. 5, 1934, p. 254.

Descriptive formula:  $S_8 = {}^a : {}^z T^{-1} : ({}_{c_p} P T^{-1})_p$       Quantic number = 8;0;0;1

Legend:

$S_8$  = The situation  
records for each of

$T^{-1}$  = annually provided  
with

${}^z T^{-1}$  = 3 years

$|_p = 2$  types of new housing  
 $\left\{ \begin{array}{l} \text{families in houses} \\ \text{families in apartments} \end{array} \right.$

${}^a : {}^z$  = with corresponding (intermittent) dates

${}_{c_p} P$  = the number of families in percentage units

Comment:

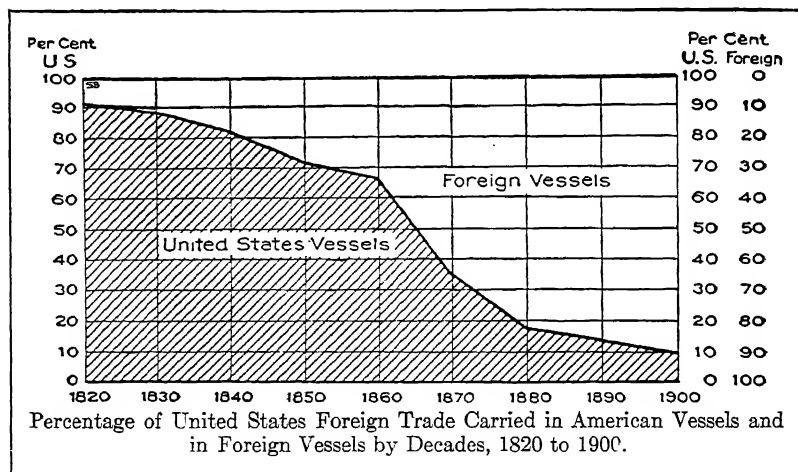
The unit of societal force here is in attribute-family-year terms, i.e., one family ( ${}_p P$ ) housed ( $I^1$ ) per year per year. In percentage units the force denoted

## S. 8 (Continued)

by "houses" ( $I_p^0$ ) and the force denoted by "apartments" ( $I_p^1$ ) are negatively correlated ( $(I_p^0, I_p^1)P = P_{p,1} = P_p$ ). House building is decelerating, apartment building is accelerating. Their sum is held constant by the use of percentage units. One P-force (in relative units) is displacing another.

$$F, \cdot F_{,,} = -1.0, \text{ where } F, = T^{-2}_{c_0}P, \text{ and } F_{,,} = T^{-2}_{c_0}P_{,,}$$

## S. 9



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 139.

Descriptive formula:  $S_9 = {}_tT^{-1} : {}_{c_0}(IT^{-1})_1$

Quantic number = 8;1;0;0

Legend:

$S_9$  = The situation  
records for each of

${}_t|$  = 8 periods

of

$T^{-1}$  = a decade each

$'|$  = since 1820

$I$  = the tonnage of foreign shipping

$T^{-1}$  = per year

${}_c|$  = in percentages of the total

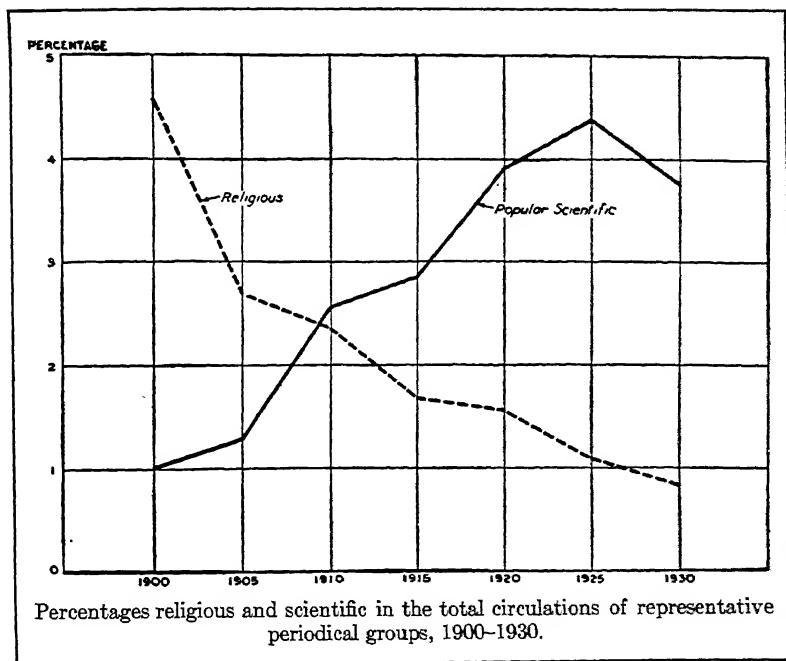
for

$|_1$  = 2 indicants—domestic and foreign vessels

Comment:

This situation illustrates an I-force which averages 1% of trade per year for the United States.

## S. 10



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, Vol. I, 1933, p. 391.

Descriptive formula:  $S_{10} = {}_tT^{-1} : \% (IT^{-1})_i$

Quantic number = 8;1;0;0

Legend:

$S_{10}$  = The situation  
records for each of

${}_tT^{-1}$  = the 6 five-year periods

' = since 1901

$IT^{-1}$  = an annual circulation  
of

$|_i$  = 2 types of periodicals

$\%$  = expressed in percents

Comment:

Two I-forces in units of percents per year per year for one plurel are illustrated in this graph. In general, by these indices, religious interest was a negative force, while scientific interest was a positive force, i.e., the former interest decelerated, the latter accelerated.

## S. 11

TREND OF RATIO BETWEEN PRODUCTION EFFORT  
AND COMMERCIAL EFFORT FROM 1850-1920

	<i>Production Effort</i>	<i>"Commercial Effort," Selling and Distribution</i>
1850.....	80.2%	19.8%
1860.....	75.1%	24.9%
1870.....	72.0%	28.0%
1880.....	67.2%	32.8%
1890.....	63.3%	36.7%
1900.....	59.9%	40.1%
1910.....	53.5%	46.5%
1920.....	49.6%	50.4%

Ref.: Chase, Stuart, *The Tragedy of Waste*, Macmillan, 1925, p. 213.

*Descriptive formula:*  $S_{11} = {}_tT^{-1} : \% (IT^{-1})_i$

*Quantic number* = 8;1;0;0

*Legend:*

$S_{11}$  = The situation

of

records for each of

${}_tT^{-1}$  = 8 decades

$|_i$  = 2 kinds  $\left\{ \begin{array}{l} \text{production} \\ \text{and} \\ \text{distribution} \end{array} \right.$

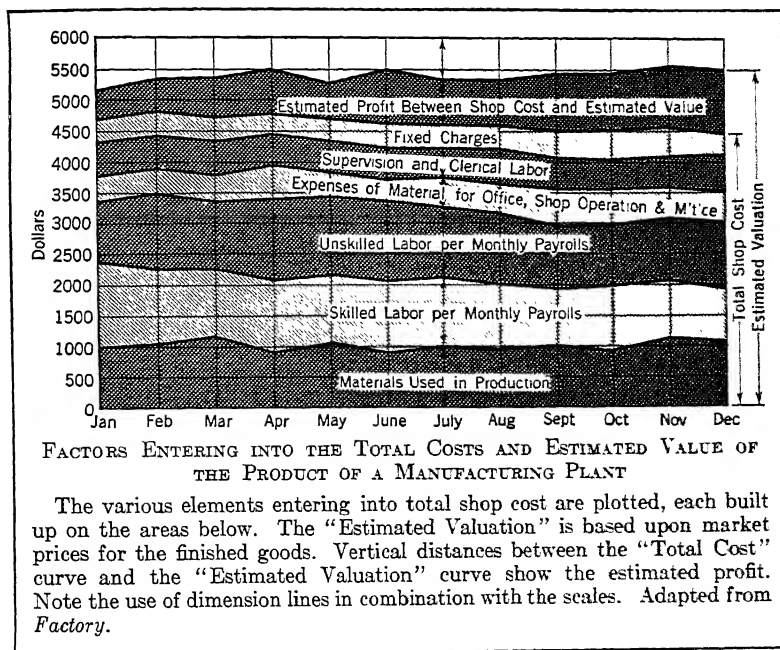
'| = beginning in 1850

$T^{-1}$  = an annual

$\%$  = expressed in % units

I = index of "effort"

## S. 12



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 147.

Descriptive formula:  $S_{12} = {}_tT^{-1} : {}_s(IT^{-1})_i$

Quantic number = 8;1;0;0

Legend:

$S_{12}$  = The situation  
records for each of

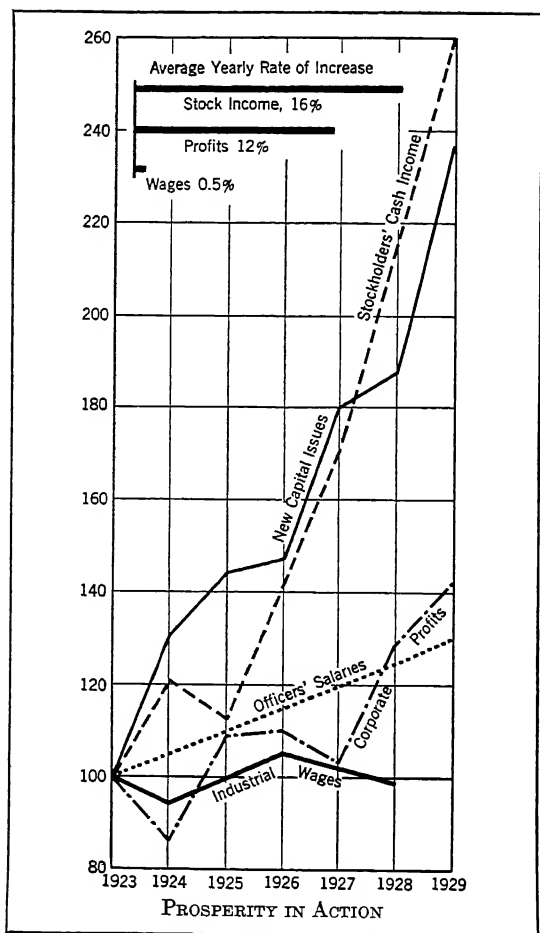
$T^{-1}$  = per month  
of each of

${}_tT^{-1}$  = 12 months

$|_i$  = seven factors of manufacturing

${}_sI$  = the cost in dollar units

S. 13



Ref.: Corey, Lewis, *The Decline of American Capitalism*, Covici Freide, 1934, p. 69.

Descriptive formula:  $S_{13} = {}_tT^{-1} : m \cdot (IT^{-1})_i$

Quantic number = 8;1;0;0

Legend:

$S_{13}$  = The situation

of

records for each of

$T^{-1}$  = dynamic (annual)

${}_tT^{-1}$  = 7 years

I = economic indices,

'| = beginning in 1923

and some of

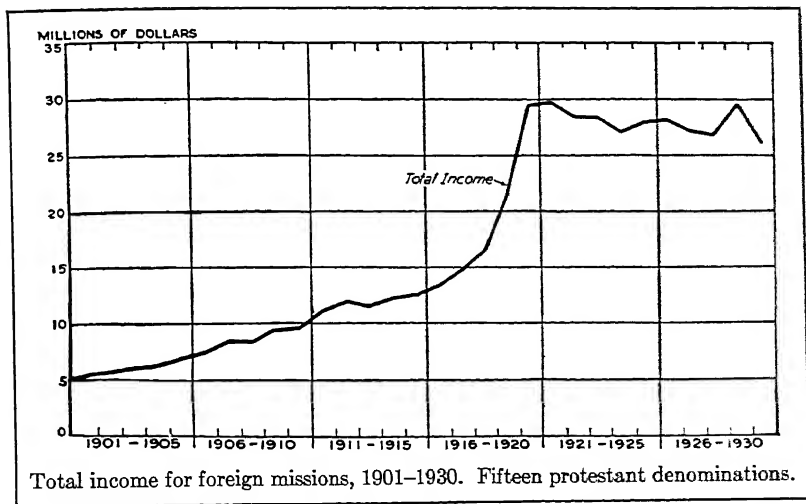
$|_i$  = 5 kinds

$m \cdot |$  = their means

Comment:

The average acceleration is given explicitly as 16% for stock income, 12% for profits, and .5% for wages. The units of I-force are here percentage units per (implied) plurel per year per year.

## S. 14



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, Vol. II, 1933, p. 1049.

Descriptive formula:  $S_{14} = \underline{P}_{\Sigma p} : \{T^{-1} : s\} (IT^{-1})$       Quantic number = 8;1;0;1

Legend:

$S_{14}$  = The situation

records for

$\underline{P}_{\Sigma p}$  = 15 Protestant denominations

for each of

$T^{-1}$  = 6 quinquennia

' = beginning in 1901

I = the gifts to missions

s = in dollars

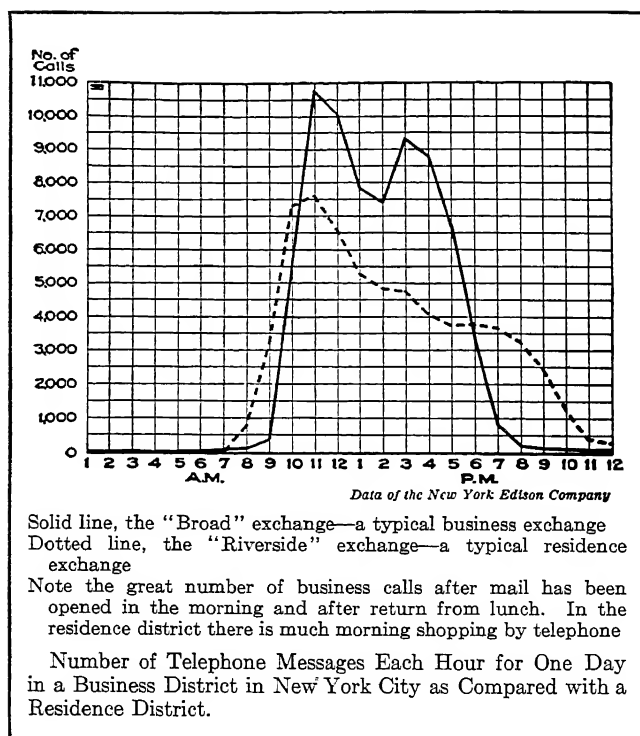
$T^{-1}$  = per year

Comment:

The straight-line trend of the first third of the century in increasing the annual income for foreign missions from 5 to almost 30 millions means an average acceleration of 5/6 of a million dollars per year per year. The straight-line trend (in common with all averages) obscures deviations from itself and does not tell of the acceleration of the war psychology and prosperity in 1918 and 1919, graphed as a steeper slope of the curve, and of the subsequent deceleration.

The curve shows the fluctuations of a societal force in units of dollars per plurel per year per year,  $I/P_p T^{-2}$ . With only one plurel (the 15 denominations together) dollars per plurel,  $I/(P^0)$ , (or  $I/\underline{P}_p$ ), is equivalent to plurel-dollars ( $P^0 \times I$ ), because multiplying and dividing by unity are numerically equivalent.

## S. 15



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 108.

Descriptive formula:  $S_{15} = \underline{P}_p : tT^{-1} : (IT^{-1})$

Quantic number = 8;1;0;1

Legend:

$S_{15}$  = The situation

for each of

records for each of

$tT^{-1}$  = the 24 hours of the day

$\underline{P}_p$  = 2 New York City telephone districts

I = a frequency of telephone calls

$T^{-1}$  = per hour

Comment:

This acceleration is an aggregation of two I-forces, as the two types of telephone calls from two telephone exchange districts imply two clientele plurels. The F-unit, i.e., unit of societal force, is here a telephone call per plurel per hour per hour. This unit is derivable from the descriptive formula above by the following series of equivalences:

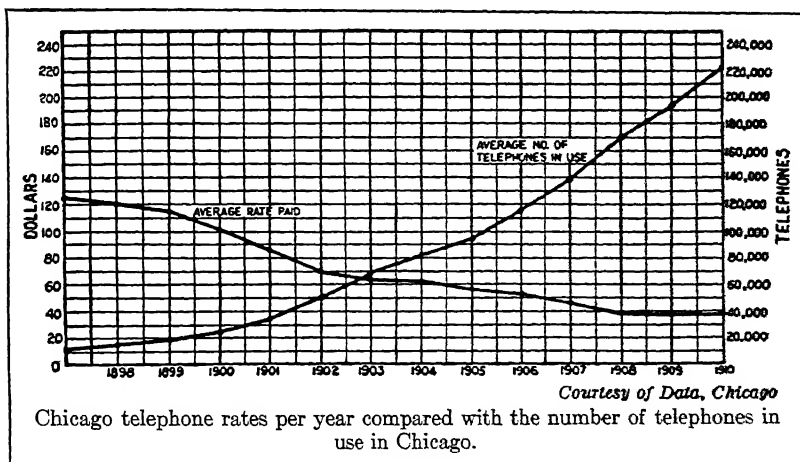
$$\underline{P}_p : (I) = (P^0)_p : (I) = (P^0)_p (I) = (I) / (P^0)_p = (I)_p$$

These symbols are verbalized as, "one plurel with a corresponding index is equivalent to a unit-plurel ( $P^0 = 1$ ) with a corresponding index, which is equiv-

## S. 15 (Continued)

alent to a unit-plurel's index, which is equivalent to an index per unit-plurel (when there is but one plurel), which is equivalent, in Brief-S notation, to an index of one plurel." When the index preceding the colon is both singular in number and qualitative, aggregating it, multiplying by it, and dividing by it have a similar result in denoting the subsequent entity as modified or qualified by that preceding quality.

## S. 16



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 136.

Descriptive formula:  $S_{16} = \underline{P}_, : {}_tT^{-1} : (I, T^{-1}, I_{,,})$       Quantic number = 8;1;0;1  
Legend:

$S_{16}$  = The situation  
records for

$\underline{P}_,$  = Chicago

for each of

${}_tT^{-1}$  = 14 years

'| = beginning in 1897

$I,$  = an index of rates

$T^{-1}$  = per year

and

$I_{,,}$  = an index of telephones in use

Comment:

This situation illustrates the tension theory:

$${}_t(PD = VE)_{r,p'} \quad \text{or} \quad {}_t[P, (I_{,,,} T^{-1}) = I_{,,} (I, T^{-1})]$$

where  $\underline{P}_,$  = the population of Chicago as the unit plurel to whom telephones were a desideratum \* where  $V = I_{,,}$  = the number of telephones, the quantity

\* Since the population data are not stated in the situation as recorded, the equations are worked out for a plurel, the city of Chicago, as the unit of population, instead of one person as the unit. The plurel unit yields the tension and force for the plurel; while the person unit yields these in per capita terms.

S. 16 (*Continued*)

of the desideratum, where  $E = I, T^{-1}$  = the tensing, measured here as an approximation, by the rate paid  $\dagger$  in dollars per year and where  $D = I,,, T^{-1}$  = the unknown average intensity of desiring of the population for the use of telephones for a year.

In this situation,  $D$  can be solved for and found as:

$D = I,,, I, / T$  = dollars per year for telephones in Chicago, a measure of the intensity of desire by this plurel for the desideratum, "telephones," for a year

For this 14-year period,  $D$  was: $\ddagger$

	$D$		$D$		$D$
1897	\$1,375,000	1901	\$3,256,000	1906	\$6,102,000
1898	2,160,000	1902	3,500,000	1907	6,580,000
1899	2,340,000	1903	4,550,000	1908	6,800,000
1900	2,500,000	1904	4,800,000	1909	7,488,000
		1905	5,415,000	1910	8,658,000

The three curves measure the average velocities of the three processes as read from the graph:

$\downarrow D$  = valuating (of telephony) = \$4,680,000 for all phones per year

$\downarrow V$  = progressing (co-operating) in respect to telephony = 16,000 new telephones per year

$\downarrow E$  = detensing in respect to telephony = \$6.40 per phone per year

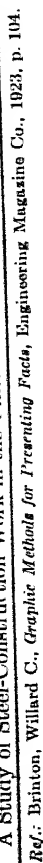
In terms of societal forces the rate of change between any two years of the "rate paid" measures the net societal force in respect to telephony in Chicago in that period.

$$F = T^{-2}IP \quad \text{in general and} \\ = T^{-1}(I, T^{-1})(P^0), \text{ in this situation, which is an } I\text{-force}$$

Thus the force is the acceleration of tensing in this plurel. The unit here, of force, is dollars—dollars per phone per year per year per plurel, i.e., the rate of change in the "rate paid" for a telephone in Chicago. The average force throughout the thirteen years was—\$6.38 $\pm$ , i.e., an average annual decrease of \$6.38 in the annual rate per phone in Chicago. This is the net force, resulting after all the unknown facilitating and resisting subforces have been vectorially summed. As usual it is directly comparable only with such other forces as are stated in the same units, or interconvertible units such as dollars into cents, or years into months.

$\dagger$  This assumes a stable price level. Inflation or deflation of the currency and similar factors make the use of price to measure economic value (which is the ratio of demand to supply) inaccurate until appropriate corrections are made. Economic value is tension,  $E$ , in this volume which is correlated with the rate paid, i.e., price, but the correlation may be imperfect depending on the stability of the price level.

$\ddagger$  These figures for  $D$  are subject to correction by a competent economist's correcting  $E$ , so that the "rate paid" will measure most consistently throughout this period the economic value.



S. 17 (*Continued*)

*Descriptive formula:*  $S_{17} = \underline{P}_r : t : {}_uT^{-1} : (IT^{-1})_i$       *Quantic number* = 8;1;0;1

*Legend:*

$S_{17}$  = The situation  
          records

'| = beginning in 1896

I = the steel construction in tons

$\underline{P}_r$  = for the United States  
          for each of

$T^{-1}$  = per month

of each of

$tT^{-1}$  = 15 years

$|_i$  = 4 kinds

$:_u|$  = subdivided into months

*Comment:*

The units ("F-units") of this societal I-force are here millions of tons per plurel per month per month,  $F = T^{-2}IP$  where:

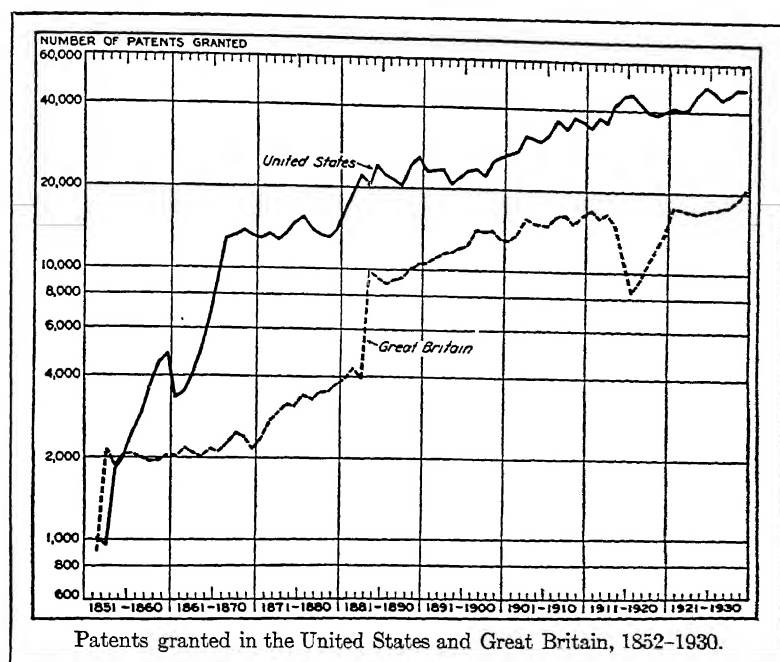
$T$  = months

$I$  = tons  $\times 10^6$

$P_r$  = the United States plurel

This force can be directly compared only with other forces which are stated in month-ton-plurel units, or multiples of them. For greater comparability the tons could be converted to a suitable index number and be compared with index numbers of diverse commodities and societal conditions.

## S. 18



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, one volume edition, 1933, p. 306.

Descriptive formula:  $S_{18} = \{T^{-1} : \underline{P}_p : (IT^{-1})$

Quantic number = 8;1;0;1

Legend:

$S_{18}$  = The situation

for each of

records for each of

$\underline{P}_p$  = 2 national plurels

$\{T^{-1}$  = 8 decades

I = an index of patents

'| = beginning in 1851

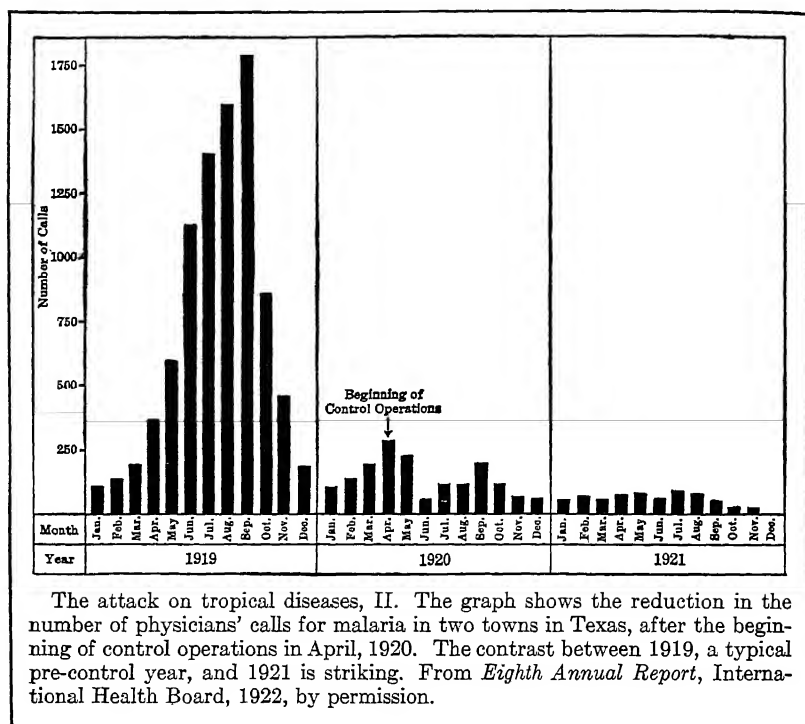
$T^{-1}$  = granted per year

Comment:

The societal force is expressed in units of one patent per plurel per year per year. Thus in Great Britain in 1881-82 it was +6000 F-units, i.e., a jump of 6000 patents, while in the United States it was several thousand negative F-units, as patenting fell off that year.

If each patent is considered as a qualitative entity rather than as an equal and interchangeable (i.e., cardinal) unit measuring the patenting process, the formula instead of  $(IT^{-1})$  would be written  $(I_{\Sigma}^0 T^{-1})$ , or  $\{I_{\Sigma}^0$  in Brief-S notation. This is the formula for the velocity of the process of dissimilarizing, Eq. 25, Ch. X, which is here the differentiating of culture through invention. The graph shows the force of invention (in each of these two countries, in the period specified, as measured by number of patents).

## S. 19



The attack on tropical diseases, II. The graph shows the reduction in the number of physicians' calls for malaria in two towns in Texas, after the beginning of control operations in April, 1920. The contrast between 1919, a typical pre-control year, and 1921 is striking. From *Eighth Annual Report*, International Health Board, 1922, by permission.

Ref.: Hankins, Frank Hamilton, *An Introduction to the Study of Society*, Macmillan, Revised edition, 1935, p. 293.

Descriptive formula:  $S_{19} = \underline{P}_{2p} : t : {}^uT^{-1} : (IT^{-1})$       Quantic number = 8;1;0;1

Legend:

$S_{19}$  = The situation

records for

$\underline{P}_{2p}$  = 2 towns in Texas (population not stated)

for each of

${}^uT^{-1}$  = 3 years

$:u|$  = subdivided by months

$'$  = beginning in 1919

$T^{-1}$  = the monthly

I = number of doctors' calls as an indicator of malaria

Comment:

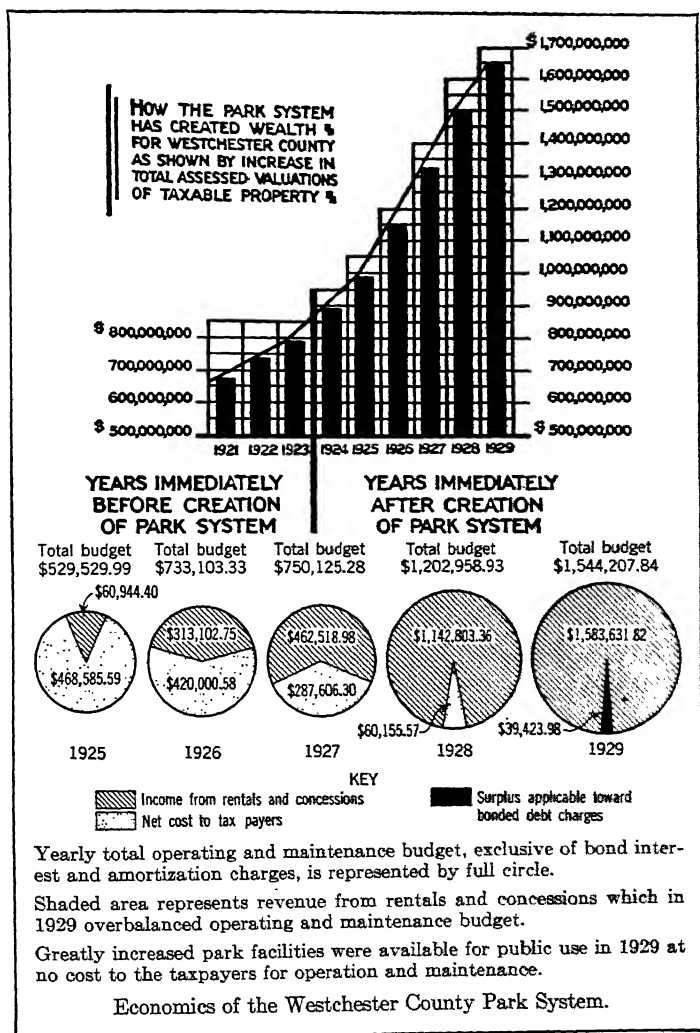
Since the population is not stated the population unit is one plurel, identified as "2 Texas towns." The unit of societal force is thus one call per month per plurel. The difference in altitude between any two adjacent columns in the graph measures the force in that month.

The compounding of forces is readily inferred. In 1919 the accelerating of malarial cases, a positive force, rose with seasonal temperature to a maximum in September and then decelerated, i.e., became a negative force, till January.

## S. 19 (Continued)

In 1920 the annual cycle started upward again, but in April the resisting force of the malarial control operations was introduced with a resultant of a negative force, i.e., decelerating (with fluctuations) for the rest of the year. In 1921 the negative control force neutralized the positive seasonal force with the result that malaria hardly rose above its winter level.

## S. 20



Ref.: Macy, Everit V., "Parks in the Modern Manner," *Survey*, Vol. LXIV, No. 7, July 1, 1930, p. 303.

## S. 20 (Continued)

*Descriptive formula:*  $S_{20} = \underline{P} : \mathbf{t} : \mathbf{t}^{-1} : (\mathbf{I}, \mathbf{I} \mathbf{T}^{-1})$     *Quantic number* = 8;1;0;1

*Legend:*

$S_{20}$  = The situation  
          records for

$\mathbf{t} |$  = with initial dates stated  
 $\mathbf{I}$ , = the tax valuations

$\underline{P}$  = Westchester County  
          for each of

and  
 $\mathbf{T}^{-1}$  = the annual

$\mathbf{t} |$  = 2 periods (before and after  
          park system)

$\mathbf{I}$  = budget  
          in

$\mathbf{t}^{-1} \mathbf{T}^{-1}$  = subdivided into years

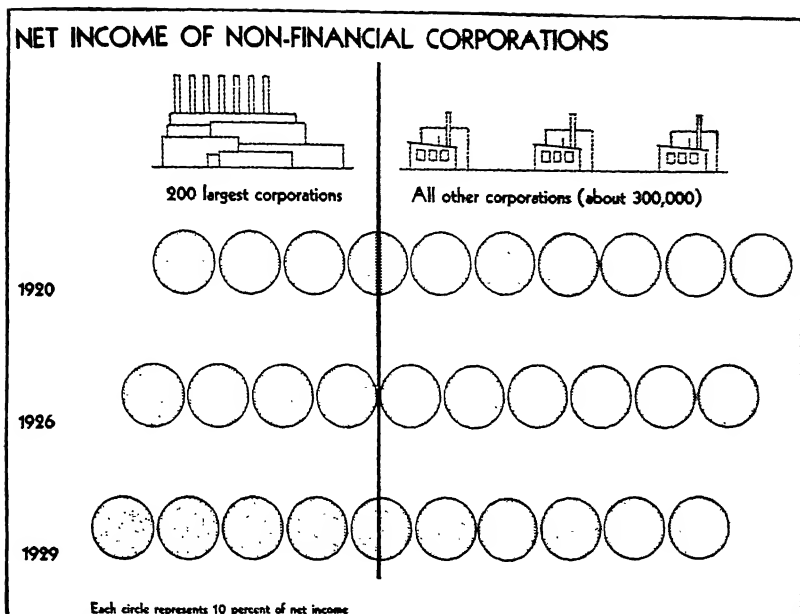
$|_2$  = 2 categories—rentals and  
          taxes

*Comment:*

In this situation one force completely displaces another. The accelerating income from rentals for this plurel more than displaces the decelerating income from taxes. Both forces are in units of dollars per year per year per plurel. Expressed in terms of percentages of the whole budget, the two forces are perfectly negatively correlated. ( $r = -1.0$ .)

Expressing these facts in vectorial terms means that the two forces can be represented by 2 vectors which are collinear but oppositely directed. Each year the negative taxation vector cuts down the taxation process, but the positive rental vector compensates and even adds enough so that the total income vector grows longer year by year.

## S. 21



Ref.: Research Bulletin of the National Education Association, Vol. XII, No. 5, Nov., 1934, p. 258.

Descriptive formula:  $S_{21} = a : z : T^{-1} : P_p : \Sigma_a : \Sigma_c (IT^{-1})$  Quantic number = 8;1;0;1  
 Legend:

$S_{21}$  = The situation  
 records for each of

$T^{-1}$  = 3 years

$a : z$  = intermittently dated, i.e.,  
 with bounding dates stated  
 for each of

$P_p$  = 2 corporation plurals  
 each composed of

$\Sigma_a$  = a stated number of corporations

$I$  = the income

$T^{-1}$  = per year

$\Sigma_c$  = in % units

*Comment:*

The nine-year trend of 1920-29 would lead to a complete monopoly of corporate income by the 200 largest corporations in about 90 years; the three-year trend of 1926-29 leads to a monopoly in about 45 years; the acceleration of the trend between 1920-26 and 1926-29, if continued, would lead to such a monopoly in about 20 years. But the fallacy in thus projecting the trend of a few years far into the future is obvious. It neglects counteracting factors which may be growing but are not yet effective, such as the depression of 1930-33, and Government regulation. To predict that the trend to concentrate wealth is likely to continue in the near future may be highly probable; but to predict that such concentration will reach a monopoly in a given period is highly uncertain.

S. 21 (*Continued*)

In calculating the amount of competing, the national corporate income as divided among 300,200 competitors in 1920 gives in round numbers,  $C_p = 1.7\%$  which rose to  $2.2\%$  in 1929. This is of the order of  $2\%$  of a national monopoly of a single corporation. If, however, the 200 largest corporations are treated as if organized into one competitor, to consider the extreme case, the percentage of maximal competing,  $C_p$ , is  $25\%$  in 1920, rising to almost  $32\%$  in 1929. Depending on how closely the 200 corporations are actually interlocked towards functioning as one, the competing lies between the limits of  $1.7\%$  and  $25\%$  in 1920, and between  $2.2\%$  and  $32\%$  in 1930. The evidence from both sets of limits agrees upon a marked increase in the direction of a monopolistic corporative State.

In the situation as recorded, there are two forces which are negatively correlated. There is the positive force gaining income for the 200 largest corporations, and the negative force whereby the second plurel of the "other corporations" lose income relatively. Each force is expressed in units of  $1\%$  of income per plurel per year per year.

## S. 22

The value of industrial production more than doubled between 1928 and 1932; the totals being 15,818 million roubles and 36,813 million respectively.

The actual figures in relation to the Five-Year Plan estimates and to 1913 are:

(In million roubles at 1926-27 prices)

	1913	1928	1932	Five-Year Plan Estimates	1932 Percentage Increase over 1928
Total Industry	10,251	15,818	36,813	36,600	132.7
Group "A"	4,290	7,024	20,486	17,400	191.6
Group "B"	5,961	8,794	16,327	19,300	85.6

Ref.: *U.S.S.R. Handbook*, Victor Gollancz Ltd., 1936, p. 137.

*Descriptive formula:*  $S_{22} = \underline{P}_p, \Sigma_p : a : zT^{-1} : (\sigma_i, IT^{-1})_1$  *Quantic number* = 8;1;0;1  
*Legend:*

$S_{22}$  = The situation

records for each of

$\underline{P}_p$  = 2 industrial plurels, A and B

$|\Sigma_p$  = and their sum

for each of

$a : z$  = intermittently spaced

$T^{-1}$  = an annual

$I$  = production (in rouble units)

$|_i$  = planned and actual

$\%|$  = with percentages also stated

${}_tT^{-1}$  = 3 years

*Comment:*

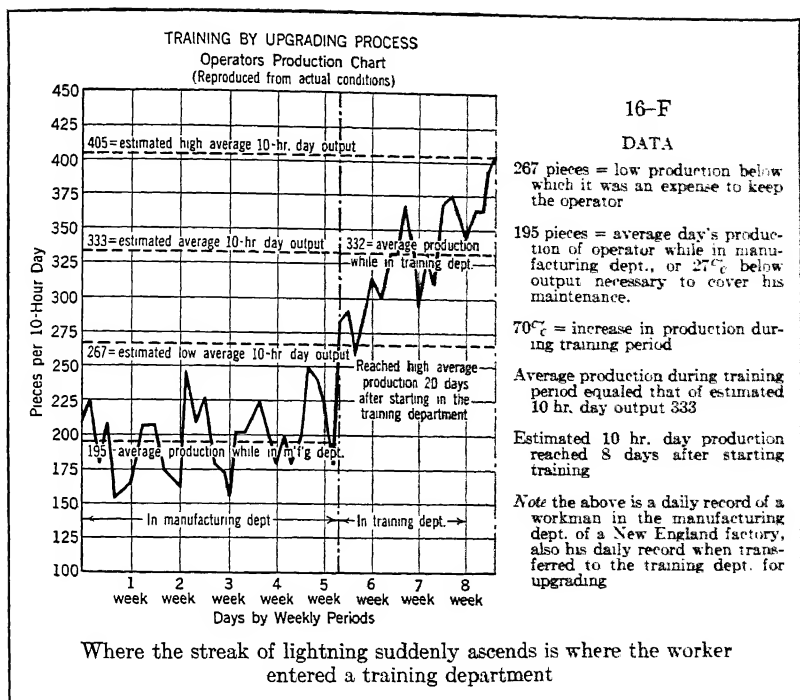
The unit of force is a rouble  $\times 10^6$  per plurel per year per year. Thus the force for total industry from 1928 to 1932 was:

36,813

15,818

$\frac{20,995}{4} = 5249$  F-units, i.e., an annual gain of 5249 million roubles of production

## S. 23



Ref.: Miles, H. E., "Making Men While We Make Materials," *Surrey*, Vol. XLIII, No. 19, March 6, 1920, p. 702.

Descriptive formula:  $S_{23} = 'P : t : u T^{-1} : i, (IT^{-1})$       Quantic number = 8;1;0;1  
Legend:

$S_{23}$  = The situation  
records

'P' = for a particular factory  
worker

for each of

$t T^{-1} = 2$  periods  $\left\{ \begin{array}{l} \text{in manufacturing} \\ \text{department} \\ \text{in training} \\ \text{department} \end{array} \right.$

$: u |$  = subdivided into 9 weeks

I = the number of pieces produced

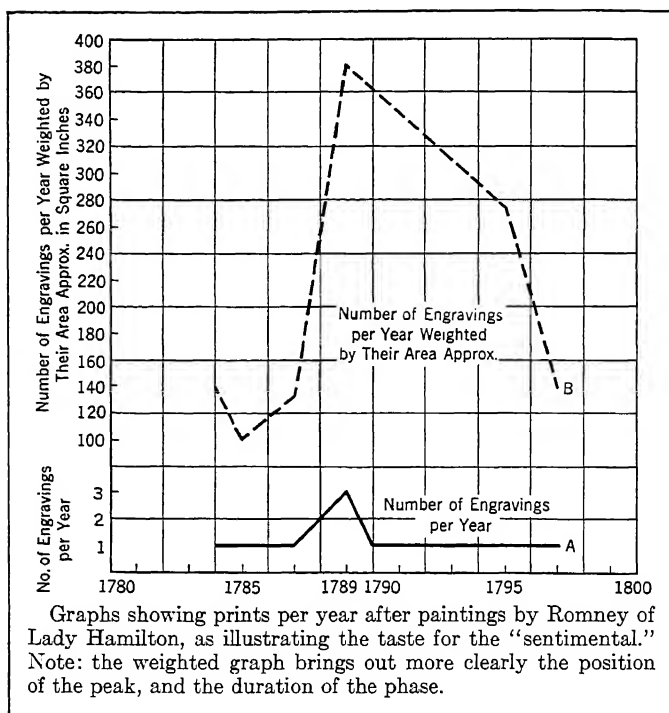
$T^{-1}$  = per day

$i, |$  = compared with 4 norm points

Comment:

The unit of force is a "piece" per person per day per day. The amount of this net force between 2 days is graphed as the difference in the vertical ordinates of those 2 successive days.

## S. 24



Ref.: Sewter, H. C., "Possibility of a Sociology of Art," *Sociological Review*, Vol. XXVII, No. 4, Oct., 1935, p. 446.

Descriptive formula:  $S_{24} = ' : ' P : \mathfrak{T}^{-1} : (I T^{-1})_1$       Quantic number = 8;1;0;1  
Legend:

$S_{24}$  = The situation

for each of

records for

$\mathfrak{T}^{-1} = 20$  years

'P = one particular person

' = beginning in 1780

: ' = in relation to another

$I_1 = 2$  indices of engraving

$T^{-1} =$  per year

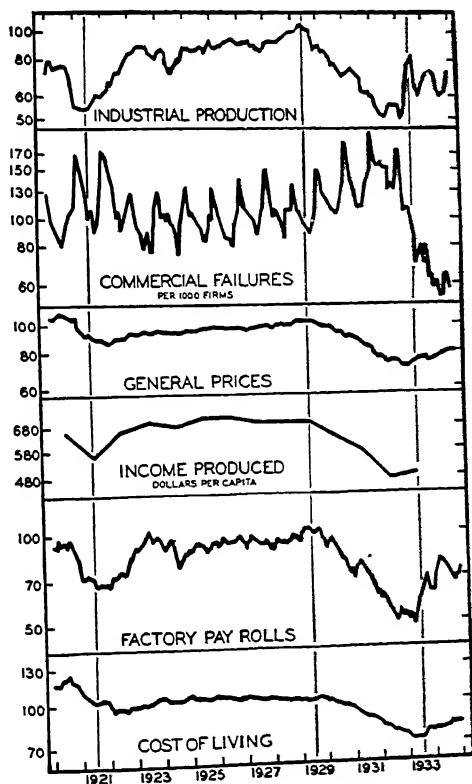
Comment:

Does this attempt to quantify tastes in art represent a societal force? It involves an acceleration ( $T^{-2}$ ), something accelerated ( $I$ ), and people ( $P$ ), which are the factors of a force,  $F = T^{-2}IP$ . The effective force, however, lies in the public's response in purchasing the engravings, not in the number of engravers ('P), nor in the number of ladies engraved (''P). The people explicitly named in the graph and symbolized in the formula by ' : ' P are not the people whose changing constitutes the effective societal force. The force is the changing taste in art in the purchasers. If the number of purchasers were stated, the "senti-

S. 24 (*Continued*)

mental" force in this situation would be measurable. As a suggestion, let  $P$ , represent the number of purchasers of a particular engraving, since an edition of an engraving is usually limited. If  $|_i$  = the number of different engravings and  $L^2$  the area of each, then  $T^{-1}(PL^2)_{\Sigma i}$  represents the ordinate of curve B when the number of editions (i.e., different engravings,  $|_i$ ) is further weighted by the number of purchasers of each edition. Then the rise and decline of this sentimental force is described by  $F = \{T^{-1} : T^{-1}(PL^2)_{\Sigma i}\}$ , the number of copies of engravings purchased, weighted by size, per year per year. The force, as denoted by the aggregative descript,  $t_i$ , is a series of forces varying from year to year as shown by the slope of the curve. Note that in this formulation the index is an area;  $L^2$  replaces  $I$  in the formula for  $F$ .

## S. 25



Ref.: Ogburn, William F., "Index of Social Trends and Their Fluctuations," *American Journal of Sociology*, Vol. XL, No. 6, May, 1935, p. 823.

## S. 25 (Continued)

*Descriptive formula:*  $S_{25} = \{T^{-1}(I)_i$

*Legend:*

$S_{25}$  = The situation  
           records for each of  
 $\{T^{-1}$  = 15 years  
       '| = beginning in 1920  
       (I) = an economic index  
           of  
       |<sub>i</sub> = 6 types

*Quantic number* = 8;1;0;1

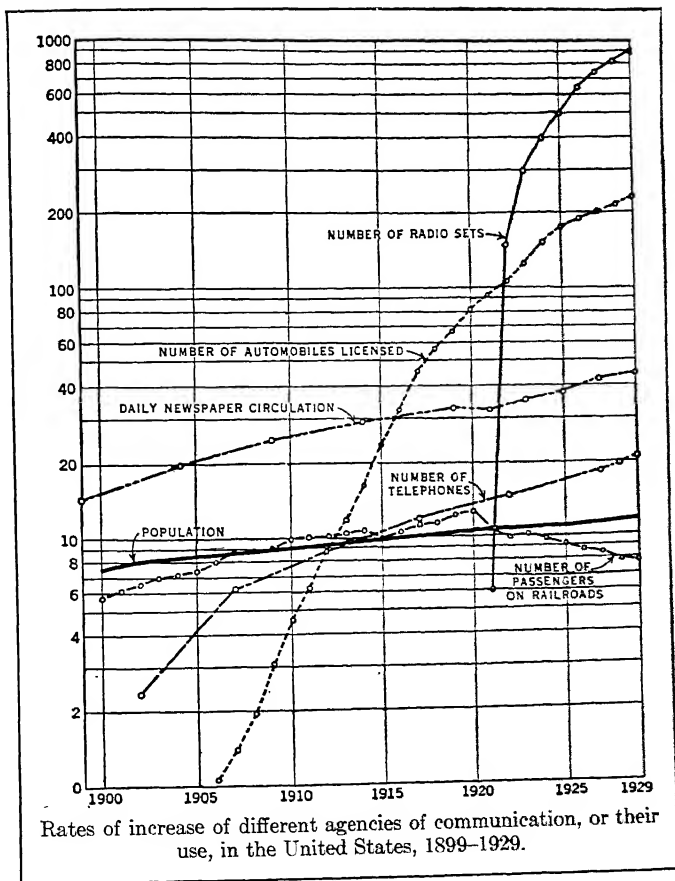
(I)<sub>i</sub> :  $\left\{ \begin{array}{l} IT^{-1} = \text{industrial production per year} \\ P\%PT^{-1} = \text{failures per 1000 per year} \\ I = \text{prices} \\ IP^{-1}T^{-1} = \text{income per capita per year} \\ IT^{-1} = \text{payroll per period} \\ IT^{-1} = \text{cost of living per period} \end{array} \right.$

*Comment:*

While all the curves show a depression-prosperity-depression cycle, the commercial failure's curve shows a seasonal cycle also.

The situation compares an aggregation of indices with varying quantics, reduced to common units of an index number with 100 as the base. As usual the highest quantic involved determines the quantic number of the situation as a whole. "Prices" represent a simple series of velocities; "income" is an "I-per-capita force" where the index is a mean, and on multiplying it by P the P cancels out, leaving the total United States income as the entity accelerated; "failures" represents a P-force; the other indices represent I-forces where the population is unity, namely the United States plurel.

## S. 26



Ref.: Cooley, Angell, and Carr, *Introductory Sociology*, Scribners, 1933, p. 165.

Descriptive formula:  $S_{26} = \underline{P}_i : {}_i^t T^{-1} \in (I)_i$   
 Legend:

$S_{26}$  = The situation  
 records in the

$\underline{P}_i$  = United States  
 for each of

${}_i^t T^{-1}$  = 30 years

Quantic number = 8;1;0;1

'| = beginning in 1899

$(I)_i$  = 6 indices of communication

$(I)_i = \begin{cases} I_i = \text{radios, autos, phones} \\ {}_i^t T^{-1} = \text{newspapers per day} \\ P = \text{population} \\ {}_i^t P T^{-1} = \text{railroad passengers} \end{cases}$

S. 26 (*Continued*)*Comment on notation:*

This aggregation of compared indices includes an I-velocity (radios), a P-velocity (population), a P-force (railroad passengers ( $I^0P$ ) per year per year), and an I-force (newspapers per day per year for the United States plural). The curves represent, therefore, 4 quantic numbers, namely, 9;1;0;0, 8;1;0;0, 9;0;0;1, 8;0;0;1, and as usual, the highest digit involved in each sector goes to make up the quantic number for the whole situation.

It may prove useful to invent some "structural formula" to show the separate quantics of the elements of this aggregated situation. Some suggestive devices are:

(1) a matrix listing the separate quantics:

T;I;L;P  
 9;1;0;0  
 8;1;0;0  
 8;0;0;1  
 9;0;0;1

(3) a braces pattern:

9; { 1;0;0  
 8; { 0;0;1

(4) an alternative digits scheme:

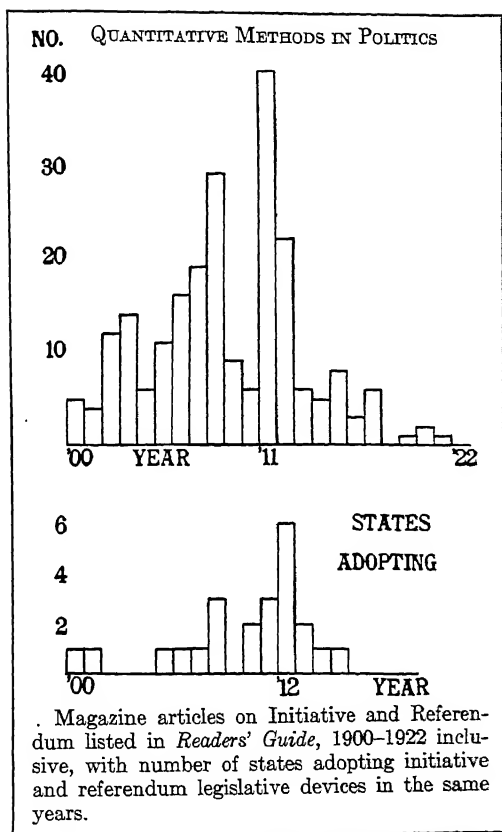
9,8: 1,0: 0:0,1

(2) a connecting line diagram:

9; ——— 1;0;0  
 \        /  
 /        \  
 8; ——— 0;0;1

Pending further exploration, the utility of such structural formulae is unknown.

## S. 27



Ref.: Rice, Stuart A., *Quantitative Methods in Politics*, Knopf, 1928, p. 248.

Descriptive formula:  $S_{27} = {}^tT^{-1} : (I, \Sigma P)T^{-1}$

Quantic number = 8;1;0;1

Legend:

$S_{27}$  = The situation

records for each of

${}^tT^{-1}$  = 22 years

' = beginning in 1900

$T^{-1}$  = the annual

$I$  = frequency of magazine articles on referenda, etc., and

$\Sigma P$  = frequency of States adopting such legislation

Comment:

The situation presents an I-force (in units of articles per year per year for the United States plurel,  $T^{-2}IP$ )<sup>0</sup> and a P-force (in units of States adopting legislation per year per year,  $T^{-2}I\Sigma P$ ). A full IP-force might be derived from a product of States and articles ( $\Sigma P I$ ), although magazine articles are a somewhat indirect indicator of the changing opinion in the States which leads up to the critical point of legislating.

## S. 28

This replacement of men by machines has only begun. One authority estimates that the average production per worker has increased about two and one-half percent per year since 1900. That figure seems much too low. Brookings Institution estimates an almost forty percent increase between 1920 and 1930 while the Report of the President's Research Committee on Social Trends states the production per man in manufacturing increased fifty-three percent between 1919 and 1929.

*Ref.: Furnas, C. C., The Next Hundred Years, Reynall and Hitchcock, 1936, p. 343.*

*Descriptive formula:*  $S_{28} = a : \mathbb{T}^{-1} : \% (IP^{-1}T^{-1})$       *Quantic number* = 8;1;0;9

*Legend:*

$S_{28}$  = The situation

records for each of

$\mathbb{T}^{-1}$  = 3 periods

$a : \mathbb{T}$  = each with its own limiting  
dates

$I$  = a production

$P^{-1}$  = per man

$T^{-1}$  = per year

$\%$  = in percentage units

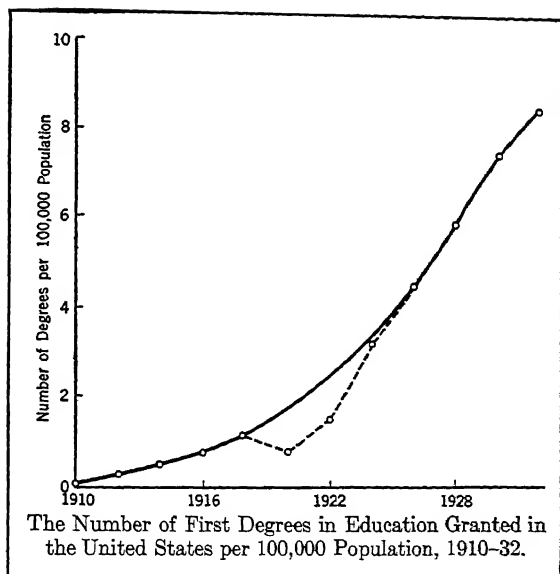
*Comment:*

This illustrates a "per capita force." It is the total force of production in the nation divided by the population to re-express it in average or per capita units. On multiplying by  $P$  the  $P$ 's cancel out, leaving an  $I$ -force ( $T^{-2}IP$ ). This is the total production of the plurel and is exactly equivalent to the per capita production ( $\sum_1^P I/P$ ) times the number of persons,  $P$ . Instead of this product,

however, the national production is actually a sum  $\sum_1^P I$ , since the productions

of individuals vary  $\sum_1^P I = (\sum_1^P I/P)P$ .

## S. 29



Ref.: Pemberton, H. Earl, "The Effect of a Social Crisis on the Curve of Diffusion," *American Sociological Review*, Vol. II, No. 1, Feb., 1937, p. 59.

Descriptive formula:  $S_{29} = {}^tT^{-1} : (IP^{-1}T^{-1})$

Quantic number = 8;1;0;9

Legend:

$S_{29}$  = The situation  
records for each of

I = the number of first degrees in  
Education

${}^tT^{-1}$  = 22 years

$T^{-1}$  = granted annually

'| = beginning in 1910

$P^{-1}$  = per 100,000 population

Comment:

This graph illustrates a per capita force where the unit is a .000,01 degree per capita per year per year. On multiplying by the population to remove the reduction to per capita units, the P cancels out, leaving the total number of degrees granted per year per year in the United States which is an I-force,  $T^{-2}IP$ .

The equation of the fitted normal probability curve is:

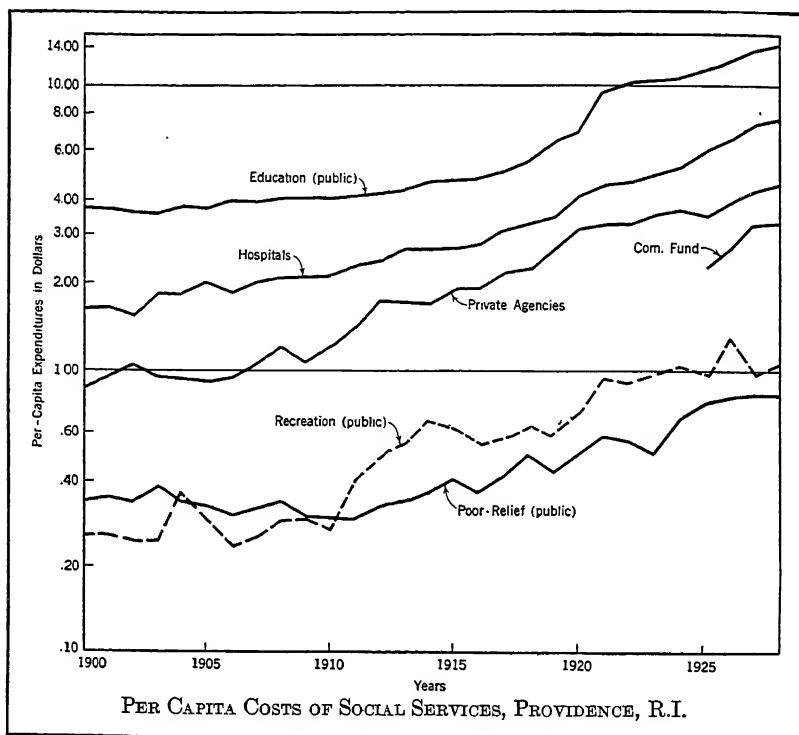
$${}_tIP^{-1} = .4IP^{-1}/\sigma(1.65^t/\sigma^2)$$

where  ${}_tIP^{-1}$  are the ordinates, the number of degrees per population in the successive years,  $IP^{-1}$  is the total number of degrees per population,  ${}^t$  is the abscissa in units of deviation from the mean date, and  $\sigma$  is the standard deviation. The remarkable finding is first, that the data follow the normal probability curve

S. 29 (*Continued*)

so closely, and second, that a major disturbing factor such as the World War is quickly compensated for after it ceases to operate. On the basis of these findings predictions of growth in the number of degrees granted annually in the near future may be made with reasonable expectation of their coming true.

## S. 30



Ref.: Phelps, Harold, and Baker, Edith, "Costs of Social Service," *Social Forces*, Vol. IX, No. 1, Oct., 1930, p. 66.

Descriptive formula:  $S_{30} = \{T^{-1} : (\$IP^{-1}T^{-1})_1$

Quantic number = 8;1;0;9

Legend:

$S_{30}$  = The situation  
records for each of

$\{T^{-1}$  = 27 years

' = beginning in 1900

$I_1$  = 6 indices of social services  
costs

$P^{-1}$  = per capita

$T^{-1}$  = per year

$\$$  = in dollar units

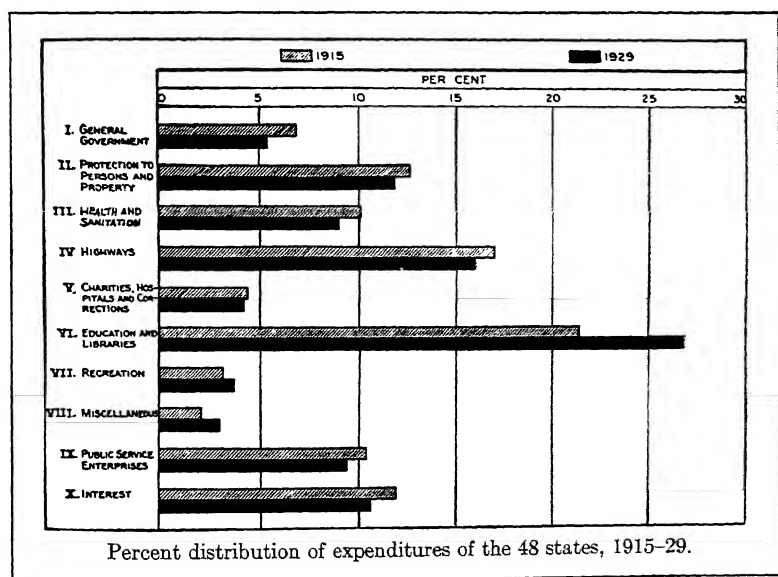
## S. 30 (Continued)

## Comment:

This is another per capita type of force which, on multiplying by the population  $P$  to eliminate reduction to per capita units, becomes an  $I$ -force ( $T^{-2}IP$ ) expressed as total social service dollars per year per year in Providence. This is actually a full societal force ("IP-force"), since  $I$  is the sum of the contributions of each person in the population, which for a distributed characteristic is equivalent to the product of a population each person of which has a constant amount of some characteristic.

For comparison of the six societal forces supporting the six social services, it is immaterial whether a per capita force, or a total IP-force, is used, since the difference between them is simply the division by a constant, the  $P$ .

## S. 31



Ref.: President's Research Committee, *Recent Social Trends*, Vol. II. McGraw-Hill, 1933, p. 1296.

Descriptive formula:  $S_{31} = {}^a : {}_t T^{-1} : \underline{P}_{\Sigma P} : \% (IT^{-1})_i$     Quantic number = 8;1;0;1  
 Legend:

$S_{31}$  = The situation

records for each of

${}_t T^{-1}$  = 2 years

${}^a : {}_t$  = with bounding dates stated

for

$\underline{P}_{\Sigma P}$  = the 48 States

$\%$  = percents of expenditures

$T^{-1}$  = annually

$_i$  = of 10 kinds

S. 31 (*Continued*)*Comment:*

This situation presents a problem in expressing the units of a societal force. Since the various subtypes of expenditures are expressed in percents of the total, the total is in effect locked to a constant 100%, regardless of whether more plurels were to come into the situation, or greater absolute expenditures in dollars were to be made. Only relative forces between the ten kinds of expenditures are presented, and these are of the I-force type, since the expenditures are for the whole plurel of 48 States ( $T^{-2}IP_{\underline{t}}$ ). Adding more States to the plurel would not increase the force proportionally, as long as it is expressed in units of percent of the total expenditure.

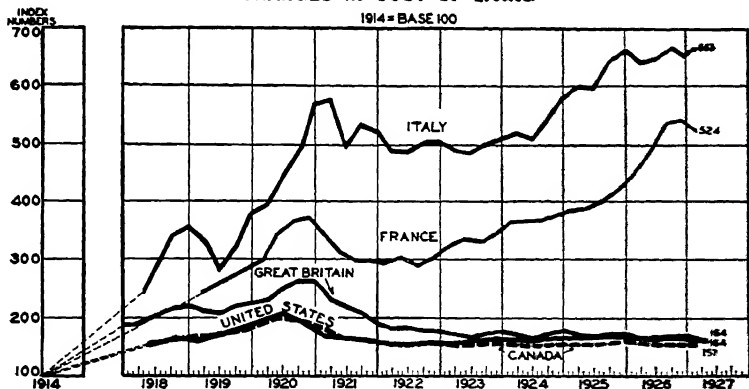
## S. 32

# COST OF LIVING IN VARIOUS COUNTRIES

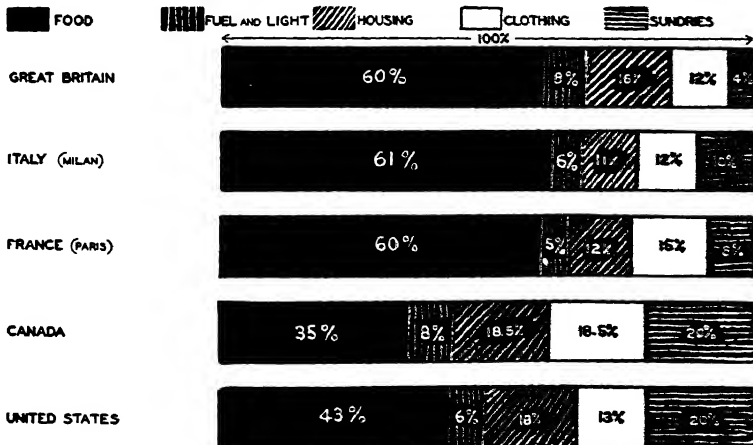
COPYRIGHTED BY  
NATIONAL INDUSTRIAL CONFERENCE BOARD, INC.  
NEW YORK CITY

## CHANGES IN COST OF LIVING

1914 = BASE 100



## MAJOR ITEMS IN FAMILY BUDGET



N. I. C. B. CHART SERVICE—CHART NO. 162  
JULY, 1927

## S. 32 (Continued)

*Descriptive formula:*  $S_{32} = \underline{P}_p : \text{t} : \text{u} \text{ } ^\text{t}T^{-1} : \% (IT^{-1})_i$     *Quantic number* = 8;1;0;1

*Legend:*

$S_{32}$  = The situation

records for each of

$\underline{P}_p$  = 5 national plurals

for every

$\text{t} : \text{u} \text{ } ^\text{t}T^{-1}$  = year and month

'| = beginning in 1914

I = a cost of living index

$T^{-1}$  = per year

|, = and compared with

|<sub>i</sub> = its 5 subdivisions

%| = in percentage units

*Comment:*

The five I-forces which produce the fluctuations in the cost of living are expressed in percentage units per year per year for one nation ( $T^{-2}\%IP$ ).

The situation illustrates the tension theory. First, as an hypothesis, let the annual cost of living be considered as the index measuring E, the (economic) tensing. Therefore, the tension theory equation becomes

$$\underline{P}_p D = VE = V(IT^{-1}) \quad (a)$$

This assumption is based on the psychology of changing costs of living where periods of rising costs are called "hard times" and unpleasant to the consumer, and falling costs mean easing of tension and fuller satisfaction (V) of desires.

Increasing costs of living ( $IT^{-2}$ ) are now identical with accelerating tension ( $ET^{-1}$ ) and this is by definition the effective economic force at work in one plural,

$$F = IT^{-2}\underline{P}_p = ET^{-1}P, \quad (b)$$

The force in these terms is an economic attensing force, i.e., it is that which accelerates this economic tension, in a plural.

As the cost of living rose in all five countries in this period the accelerated tension meant one of two things to the consuming public—either decrease of goods and services consumed, ( $-_tV$  = "regressing"), or increased desire ( $+_tD$  = "evaluating") for more monetary income to exchange for the accustomed goods and services of a previous standard of living. The tension theory thus analyzes, or expresses, the psychological as well as economic factors involved in a changing index of living costs.

Next in the family budgets the relative shares of the five kinds of items may be as the relative amounts of each of five desiderata ( $V_v$ ) whose sum is called 100% as a convenient unit:

$$V_1 + V_2 + V_3 + V_4 + V_5 = V_{\Sigma v} = 100\% \quad (c)$$

The tensions in each family for each kind of desideratum tend to be equal:

$$E_1 = E_2 = E_3 = E_4 = E_5$$

$$\text{or } \frac{D_1}{V_1} = \frac{D_2}{V_2} = \frac{D_3}{V_3} = \frac{D_4}{V_4} = \frac{D_5}{V_5} \quad (d) \quad (\underline{P}_p, \text{ being a constant, can be canceled out})$$

for if not, one kind of item will have (a) a smaller ratio, or (b) a larger ratio than the others. If the ratio is smaller it means that that item is less desired (D) relative to the dollars (V) assigned to it, and the family will tend to reassign some of the dollars of that item to other items that are relatively more desired. If any ratio is larger than the others, it has fewer dollars in the denominator relative to the family's desire for it in the numerator than is the case for other items, so that the family will tend to take dollars from other items and assign

S. 32 (*Continued*)

them to this item, until the tension for it is roughly equalized with tensions for each of the others.

By interposing ratios in pairs such as :

$$D_1/D_2 = V_1/V_2 \quad (e)$$

it is found that the ratio of the desires for any two desiderata (i.e., any two of the 5 kinds of items) is equal to the ratio of the amounts of those two desiderata. To obtain a unit for expressing these unknown intensities of desire,  $D$ , let their sum be 100%. This is the total intensity of desire for an economic living and uses a unit comparable to the percentage units in which the desiderata are measured.

$$D_1 + D_2 + D_3 + D_4 + D_5 = D_{\Sigma} = 100\% \quad (f)$$

Solving among these equations gives a relative intensity of desire for each kind of desideratum exactly equal to the relative amount of that desideratum :

$$D_1 = V_1, D_2 = V_2, D_3 = V_3, D_4 = V_4, D_5 = V_5 \quad (g)$$

and so,

$$E_1 = E_2 = E_3 = E_4 = E_5 = 1 \quad (h)$$

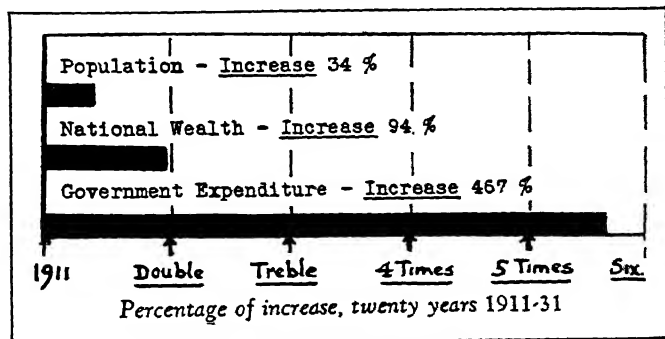
This result flows from the assumption of equal tensions and the expressions of both the desiderata and the desires in comparable units as percents of their sum. The assumption of equal tensions will be somewhat in error in individual families where a commitment, such as a house rented for a long period, may prevent prompt readjustments between the five kinds of desiderata as the family's desires change; but on the average for the nation, as in the graph above, these errors average out, and the averaged tensions for the five kinds of desiderata probably are very nearly equal, under the existing economic conditions.

Note that instead of considering the percentages for each kind of desideratum as measuring the 5 relative amounts,  $V$ , of those desiderata attained by families, these percentages could be considered as measuring the 5 relative intensities of desire,  $D$ , and the result would be the same.

This situation illustrates the solution of the matrix equation composed of 5 simple interdependent equations for the five desiderata,  $v$ , all repeated for each of the five nations,  $|_v$ , making 25 equations for one year :

$${}_v(PD = V_v E)_p \quad (i)$$

## S. 33



Ref.: Gairt, John Palmer, "Very Dear Uncle Sam," *Survey*, Vol. LXVII, No. 9, Feb. 1, 1932, p. 493.

## S. 33 (Continued)

Descriptive formula:  $S_{33} = ,T^{-1} : \% (P, I, (IT^{-1}))$ 

Quantic number = 8;1;0;1

Legend:

 $S_{33}$  = The situation  
recordsI = national wealth,  
and $,T^{-1}$  = for the 20-year period 1911-  
31

I = government expenditure

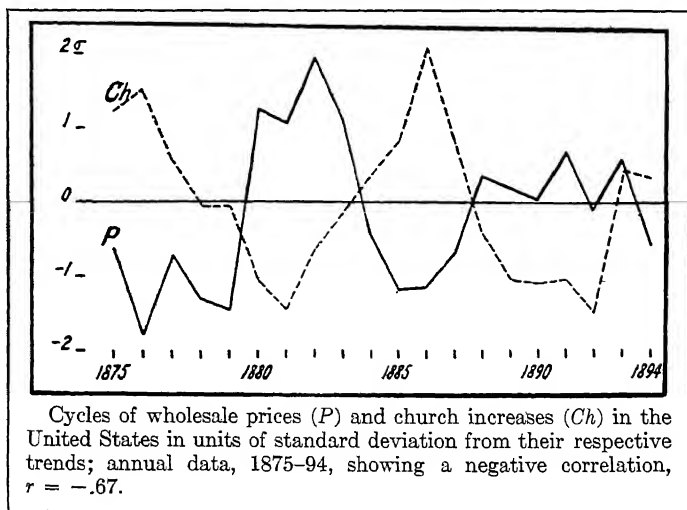
 $\%$  = the percentage increase in $T^{-1}$  = per year

P = the U.S. population,

## Comment:

A superficial application of the formula for a societal force,  $F = T^{-2}IP$ , might result in multiplying the change of the annual expenditure by the population. More penetrating analysis, however, suggests this to be a less meaningful unit than to consider the expenditure per year per year as the force (I-force). The population is implicit in the total government expenditure. The total expenditure in reality is the sum of the expenditures proratable to each citizen. As prorating each citizen's exact and differential share is difficult, the average, or per capita, expenditure serves as an approximation valid for the whole plurel, although inaccurate in showing individual deviations. The per capita share times the population is the true IP product, the acceleration of which is the societal force of government spending; but this product is simply the total expenditure. To multiply it again by the population is to take a characteristic of the persons of a population and multiply it by the square of the population—which produces a pointless result.

## S. 34



Ref.: Reinhardt, J., and Davis, G., *Principles and Methods of Sociology*, Prentice-Hall, 1932, p. 169.

S. 34 (*Continued*)

*Descriptive formula:*  $S_{34} = \underline{P}_, : \{T^{-1} : I, \bullet (IT^{-1})_{,,}$       *Quantic number* = 8,2;0;1

*Legend:*

$S_{34}$  = The situation

records for

$\underline{P}_,$  = the United States

for each of

$I,$  = an indicator of prices

$\bullet$  = correlated with

$(I)_{,,}$  = an index of church increase

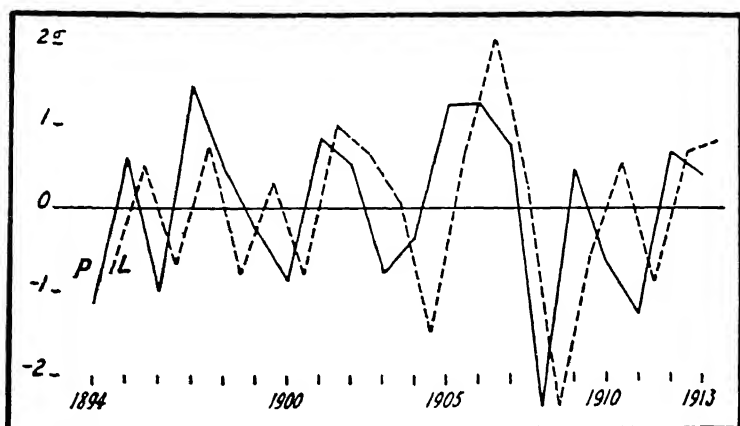
$T^{-1}$  = per year

$\{T^{-1}$  = 20 years, beginning in 1875

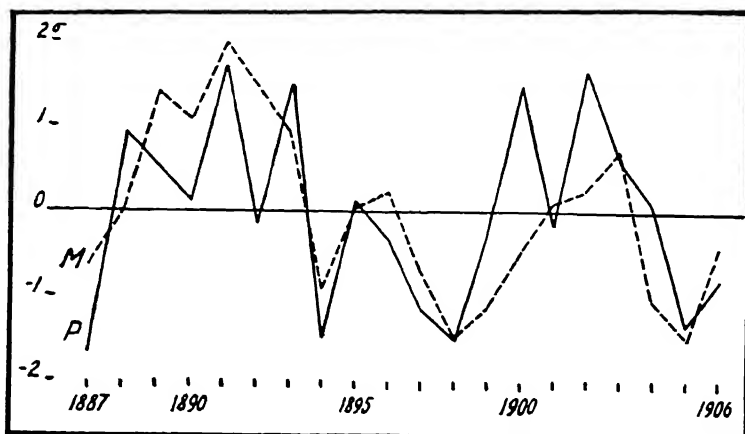
*Comment:*

The correlation of two time series illustrates the process of "recovering" (Brief-S =  $\cdot(IJ_t)$ ). The recovering is between an index of religion and of economic adversity. This correlation plus antecedence of the price changes are the two necessary (but not sufficient) conditions for proving economic adversity a partial cause of religious interest, in this situation, to the extent of a coefficient of determination of 45% (= .67<sup>2</sup>). (See "Causation" in Chapters VI and XI.)

## S. 35



Cycles of wholesale prices ( $P$ ) and per capita consumption of liquor ( $L$ ) in the United States, in units of standard deviation from their respective trends; annual data, 1894–1913, showing a positive correlation,  $r = .78$ .



Cycles of wholesale prices in the United States ( $P$ ) and the marriage rate for the United States ( $M$ ) in units of standard deviation from their respective trends; annual data, 1887–1906, showing a positive correlation,  $r = .67$ .

S. 35 (*Continued*)

*Descriptive formula:*  $S_{35} = {}^{a,z}_t T^{-1} : {}_{\sigma} I, \bullet ({}_{\sigma} I,, P^{-1} T^{-1}, {}_{\sigma\%} P T^{-1})$

*Quantic number* = 8;2;0;1

*Legend:*

$S_{35}$  = The situation

records during

${}_t T^{-1}$  = 27 years

${}^{a,z}_t$  | = in periods with stated limits

${}_{\sigma} I,$  = an index of prices (in  $\sigma$  units)

$\bullet$  = correlated with

${}_{\sigma} I,,$  = an index of liquor consumption

$P^{-1} T^{-1}$  = per capita annually

$,$  = and with

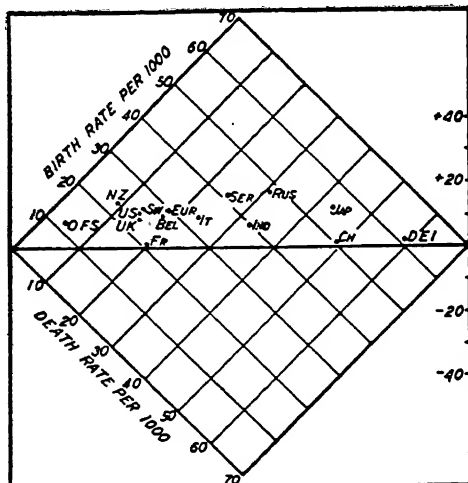
$T^{-1}$  = the annual

${}_{\sigma\%} P$  = marriage rate

*Comment:*

In this correlation of time series ("recovarying"), changes of a static index are correlated with changes in each of 2 dynamic indices. A velocity is correlated with each of 2 forces—a per capita subtype of an I-force and a P-force (newlyweds per population per year per year). A marked correlation of both liquor consumption and marrying with economic prosperity is revealed.

## S. 36



Estimates of pre-war annual birth and death rates, per 1000 population, in certain countries (see Table 22). The plus and minus scales to the right indicate the annual rate of increase or decrease per 1000 population. Countries to the left have low rates, and countries to the right have high rates. With advancing civilization there appears to be a tendency for a country to move from low right to higher middle and to lower left.

Country	1870		1910	
	Births	Deaths	Births	Deaths
Orange Free State.....	19	8	12	4
New Zealand.....	39	12	23	9
United States.....	39	21	25	14
United Kingdom.....	35	22	24	15
Sweden.....	30	19	26	14
Europe.....	37	26	30	18
Belgium.....	30	22	28	18
France.....	26	24	21	20
Italy.....	38	32	33	24
Serbia.....	44	33	41	25
Russia.....	51	37	48	31
India.....	..	..	40	33
Japan.....	..	..	55	43
China.....	..	..	50	49
Dutch East Indies.....	..	..	61	59

Estimated trends of pre-war annual birth and death rates per 1000 population in specified countries at approximately the dates 1870 and 1910.

S. 36 (*Continued*)

*Descriptive formula:*  $S_{36} = ,T^{-1} : ,,T^{-1} : i(\%P) :: i(\%P) : \underline{P}_p$   
*Quantic number* = 8;2;0;1

*Legend:*

$S_{36}$ = The situation	$i $ = in 7 class-intervals
records	$::$ = cross-classified with
$,T^{-1}$ = for the 40-year	$\%P$ = the index of the death rate
period 1870-1910	$j $ = in 7 class-intervals
$,,T^{-1}$ = the annual	$:$ = with the corresponding
$\%P$ = index of the birth rate	$\underline{P}_p$ = frequency of the 15 nations

*Comment:*

The situation illustrates acceleration of the process of recodispersing, the change of a correlation of two dynamic indices. The correlation of births and deaths in 1870 is .72, and 1910 (in the same sample of 11 nations) is .93, representing an annual velocity of change in this pattern of 5 r points per decade.

*Comment on Notation:*

By Rule 8, Appendix II, the two correlated population indices are classed in the quantic number as a correlation by the indicatory digit of 2.

## S. 37

RELATIVE IMPORTANCE OF SYRIA'S TRADE WITH NEIGHBORING COUNTRIES \*  
(In percentage)

Year	Turkey					Egypt					Palestine					Iraq					Transjordan				
	Imports from †	Transit from	Exports to ‡	Re-exports to	Transit to	Imports from †	Transit from	Exports to ‡	Re-exports to	Transit to	Imports from †	Transit from	Exports to ‡	Re-exports to	Transit to	Imports from †	Transit from	Exports to ‡	Re-exports to	Transit to	Imports from †	Transit from	Exports to ‡	Re-exports to	Transit to
1930	—	—	1.7	12.0	17.0	—	—	22.5	8.0	7.5	—	—	15.0	36.0	6.4	—	—	1.8	17.0	—	—	—	—	3.7	7.7
1931	8.0	8.0	2.0	12.0	6.7	3.0	—	12.0	4.0	4.8	2.0	—	25.0	38.0	4.8	5.8	3.6	3.0	22.0	10.0	—	—	4.5	8.0	
1932	6.7	5.5	3.0	14.0	7.7	3.0	1.0	6.5	2.0	3.2	2.5	2.0	32.0	32.0	5.0	3.2	2.3	6.0	30.0	24.0	—	—	4.0	9.0	
1933	8.0	6.5	1.9	4.0	5.2	3.3	—	4.3	1.9	3.4	2.3	—	41.5	46.0	5.5	2.0	3.0	7.9	26.0	20.2	—	—	4.0	—	

\* Derived from *Statistiques Générales*, 1930-33.

† Imports for local consumption.

‡ Exports of local production.

Ref.: Himadeh, Said B., *Economic Organization of Syria*, American Press, Beirut, 1936, p. 248.

S. 37 (*Continued*)

*Descriptive formula:*  $S_{37} = \underline{P}_i :: \underline{P}_a : \text{'}T^{-1} : \text{'}\% (IT^{-1})_i$  *Quantic number* = 8;1;0;2

*Legend:*

$S_{37}$  = The situation  
records for

$\underline{P}_i$  = Syria  
:: = cross-classified with each of

$\underline{P}_a$  = 5 neighboring countries  
for each of

$\text{'}T^{-1}$  = 4 years

$\text{'}$  = beginning in 1930

$I$  = the trade

$T^{-1}$  = per year

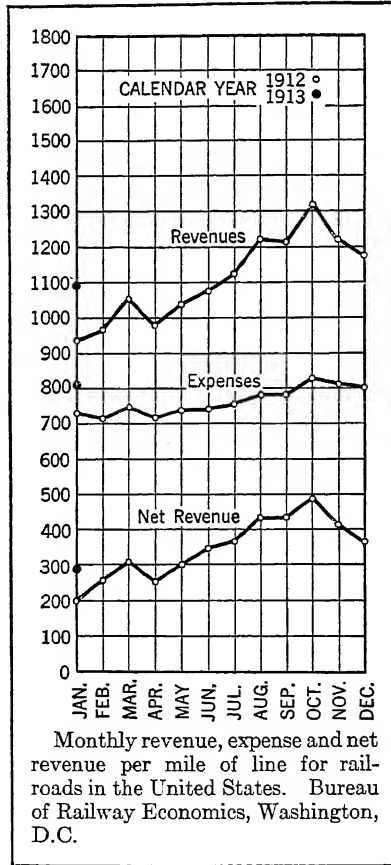
$\%$  = in percentage units  
in

$|_i$  = 5 categories (transit, re-export, etc.)

*Comment:*

This situation combines two-way interaction with acceleration defining a force. This combination of a societal force and interaction of plurels is the definition of societal control in this volume—the accelerating of change in one plurel by another plurel,  $T^{-2}IP^2$ . Each country's changes of annual imports change the exports of the other country. The unit of force is one percent of trade per year per year per plurel.

S. 38



Ref.: Brinton, Willard C, *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 257.

Descriptive formula:  $S_{38} = \underline{P}_i : \mathbb{T}^{-1} : (IL^{-1}\mathbb{T}^{-1})_i$       Quantic number = 8;1;9;1  
 . Legend:

$S_{38}$  = The situation

records for

$\underline{P}_i$  = the larger U.S. railroad companies

for each of

$\mathbb{T}^{-1}$  = 12 months

'| = beginning January, 1912

I = the dollars

$L^{-1}\mathbb{T}^{-1}$  = per mile per year in

$|_i$  = 3 classes—Revenues, Expenses, and Net Revenue

## S. 38 (Continued)

*Comment:*

The fundamental equation of accounting in all business conducted for profit is represented by the ordinate in the graph, and the series of ordinates presents the equation expanded in the time dimension into a matrix equation.

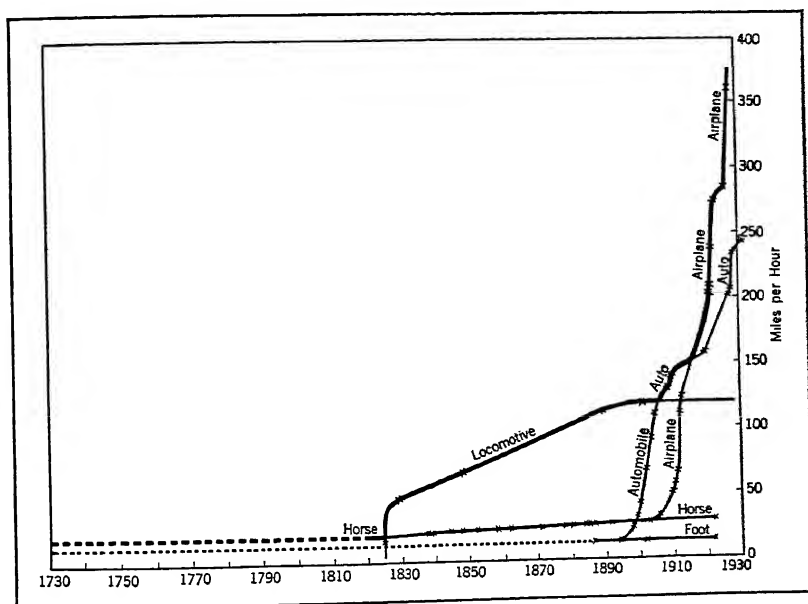
$$\text{Income, } I, T^{-1} - \text{Expenses, } I, T^{-1} = \text{Profit or loss, } I, T^{-1}$$

which reduced to mileage units for each of 12 months becomes:

$${}_tT^{-1} : (IT^{-1},, ,,)L^{-1}$$

In terms of forces, the two independent I-forces,  $F$ , and  $F,,$ , and their resultant difference,  $F,,,$ , are expressed in units of dollars per mile per month per month per plurel, where the plurel is the railroad companies specified in the caption.  $F = T^{-2}(IL^{-1})P_{\cdot}$ .

## S. 39



Ref.: Hart, Hornell, *The Technique of Social Progress*, Henry Holt, 1931, p. 76.

Descriptive formula:  $S_{39} = {}_tT^{-1} : (LT^{-1})_i$

Legend:

$S_{39}$  = The situation  
records for each of

${}_tT^{-1}$  = 20 decades

'|' = beginning in 1730

Quantic number = 8;0;1;0

$L$  = the number of miles  
 $T^{-1}$  = per hour

by each of

$|_i$  = 5 kinds of human travel

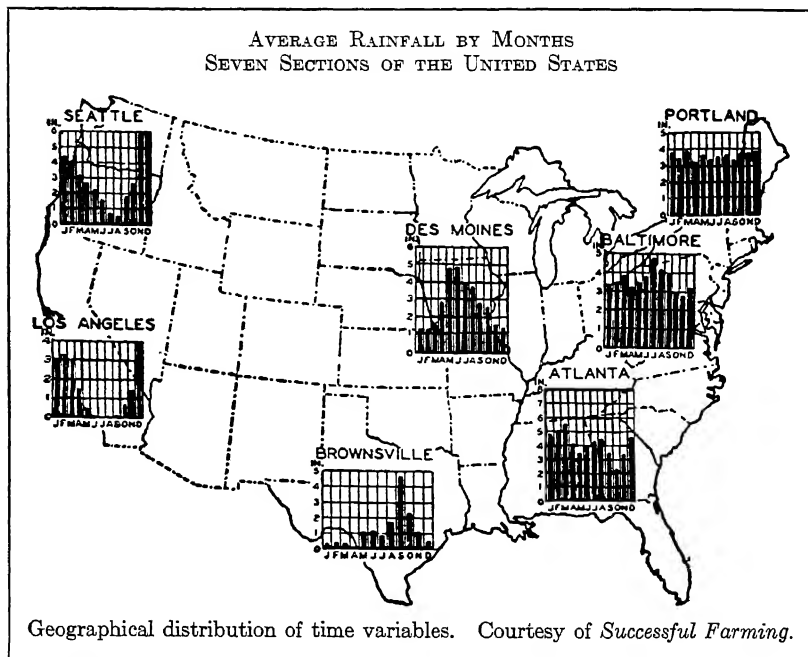
## S. 39 (Continued)

*Comment:*

One of the simplest forms of societal acceleration is this speeding up of speed of travel, the velocity of change over the decades of the hourly velocity of travel. Physical acceleration and societal acceleration are here equivalent.

By making more explicit the plurel "human travelers,"  $P$ , to whom this situation refers, since all these speeding objects are vehicles for carrying people, a societal I-force of traveling can be inferred in units of miles per hour per decade per vehicular plurel.  $F = LT^{-1}T^{-1}P$ . (Note again that for *one* plurel, as for *one* attribute, multiplying by it or dividing by it is numerically equivalent, so that "plurel-miles" is equivalent to "miles per plurel,"  $LP = L/P$ . (As most I-forces are for one plurel this principle should be remembered when expressing their units in words.)

## S. 40



## S. 40 (Continued)

Descriptive formula:  $S_{40} = {}^1L_{,1}^2 : {}^tT^{-1} : (IT^{-1})$

Quantic number = 8;1;2;0

Legend:

$S_{40}$  = The situation

for each of

records for each of

${}^tT^{-1} = 12$  months

${}^1L^2 = 7$  cities

the

in

$I$  = rainfall

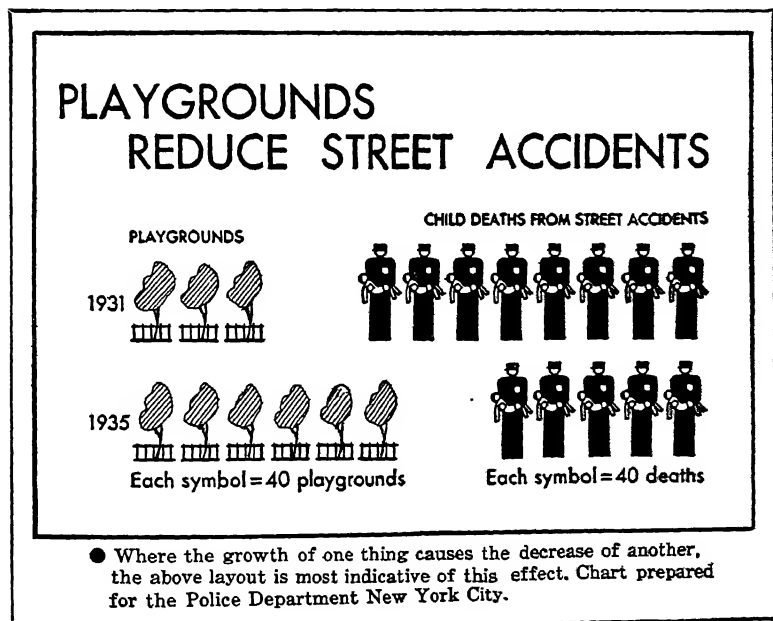
$_{,1}$  = the 48 mapped States

$T^{-1}$  = per month

Comment:

This situation is best interpreted as a simple acceleration of a physical characteristic of the human environment. The interpretation could be stretched, however, to consider it a societal I-force by inferring the people in each city and calling the units inches of rainfall per month per month per each of 7 plurels. ( $F = T^{-2}LP$ )<sub>p</sub>

## S. 41



S. 41 (*Continued*)

*Descriptive formula:*  $S_{41} = ,T^{-1} : ({}_1L^2, PT^{-1})$       *Quantic number* = 8;0;2;1

*Legend:*

$S_{41}$  = The situation

, = compared with

records for

P = children killed by accidents

$,T^{-1}$  = a four-year period 1931-35

$T^{-1}$  = per year

${}_1L^2$  = the number of playground  
areas

*Comment:*

Causation is suggested but not proven by the P-force situation as recorded. An association of a prior condition (playgrounds) and a subsequent condition (accidents) under highly urban conditions (New York City) is shown, but without enough further evidence of the frequency of such association to reliably establish the probability of it. The evidence for causation is "necessary but insufficient."

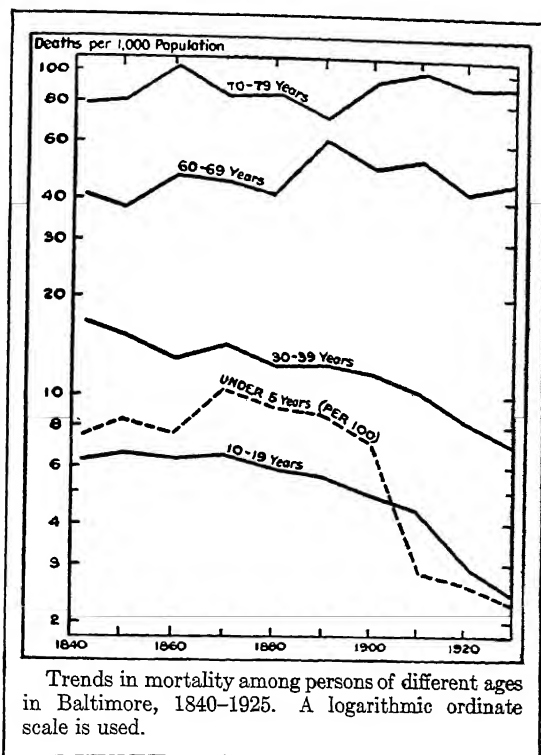
*Comment on notation:*

Note that the population index, P, is as usual a condensed implicit product of one or more attributes,  $I^0$ , and a "pure" P, thus:

$$I^0 \times I^0_{,,} \times I^0_{,,,} \times P'_0 = P_{\times \times \times \times} = \text{accidentally - killed - child - persons} =$$

$P,$	$= P$	by Eq. 4, Ch. IV
children killed by accidents	the same with singular class script understood	

S. 42



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, Vol. I, 1933, p. 607.

Descriptive formula:  $S_{42} = {}_tT^{-1} : {}^a : {}^z{}_uT^{+1} : ({}_uP T^{-1})$  Quantic number = 81,0,0;1

Legend:

$S_{42}$  = The situation  
records for each of  
 ${}_tT^{-1}$  = 9 decades  
'| = beginning in 1840

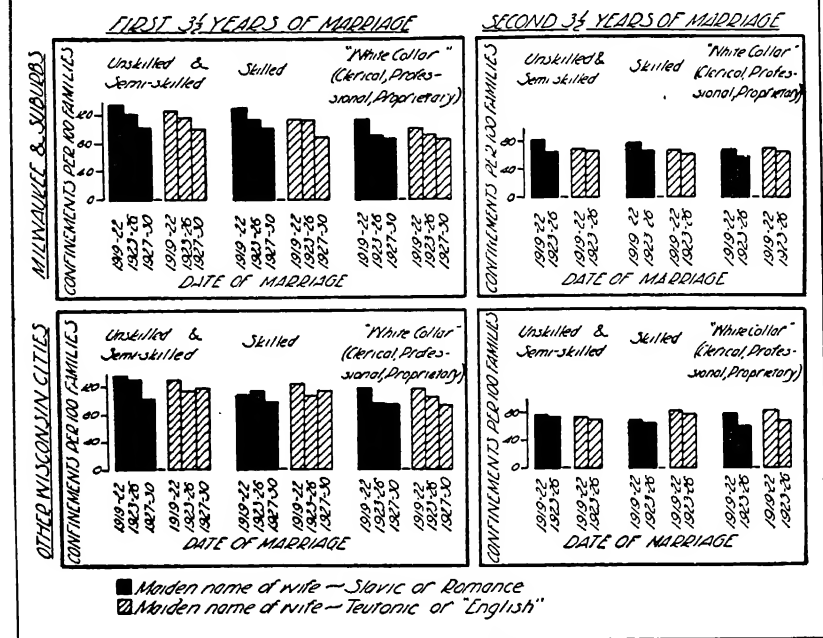
and for each of  
 ${}_uT^{+1}$  = 5 10-year age groups  
 ${}^a : {}^z$  = with varying limits  
 $T^{-1}$  = the annual  
 ${}_uP$  = death rates

Comment:

The societal force of mortality is here of the durational P-force type, i.e., it is a deceleration of depopulating classified by ages. The unit is one person dying per 1000 per year per year.  $F = T^{-2}I^0{}_uP$ .

S. 43

CATHOLICS ONLY—TRENDS IN CONFINEMENTS PER 100 FAMILIES ACCORDING TO LINGUISTIC GROUP OF WIFE'S MAIDEN NAME, BY OCCUPATION, RESIDENCE, AND INTERVAL AFTER MARRIAGE  
Rates standardized for age of wife



Ref.: Stouffer, Samuel A., "Trends in the Fertility of Catholics and Non-Catholics," American Journal of Sociology, Vol. XLI, No. 2, Sept., 1935, p. 146.

Descriptive formula:  $S_{43} = {}^tT^{-1} : {}^uT^{+1} : ({}^pPT^{-1})_p : q : r$

Quantic number = 81;0;0;1

Legend:

$S_{43}$  = The situation  
records for each of

${}^tT^{-1}$  = 3 4-year periods

${}^1$  = beginning in 1919

for each of

${}^uT^{+1}$  = 2 durations { 1st 3.5 years  
of marriage { 2nd " "

${}^pP$  = the ratio of confinements per  
100 families

$T^{-1}$  = per 3.5-year period  
classified into

${}_p$  = 2 regional plurels  
subclassified into

${}_q$  = 3 occupational plurels  
and further  
subdivided by

${}_r$  = 2 racial name-plurels

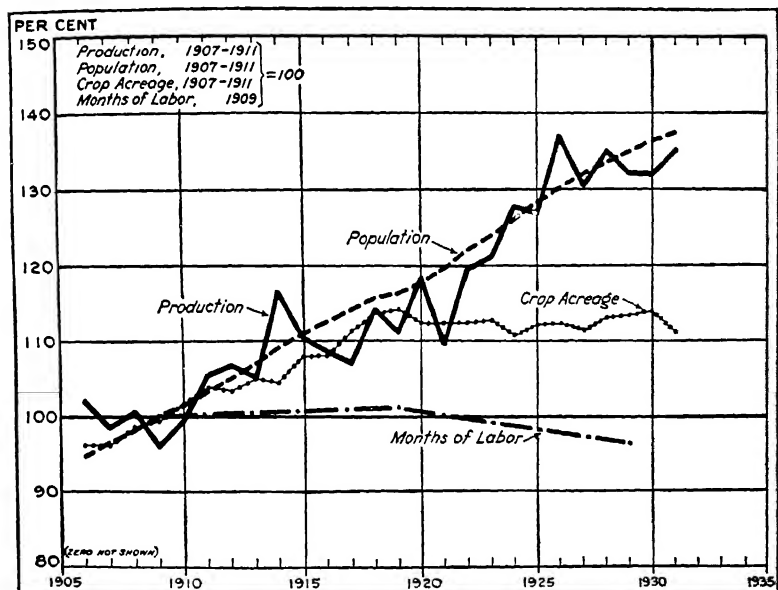
Comment:

The force here is a negative P-force, the deceleration of the adpopulating process, classified so as to isolate conditioning factors of the duration of mar-

## S. 43 (Continued)

riage, race, region, and occupation. The classifying is intended to determine whether the birth rates of Catholics are declining when freed of other irrelevant but obscuring variables.

## S. 44



AGRICULTURAL PRODUCTION, NATIONAL POPULATION, CROP LAND, AND FARM LABOR. PERCENTAGE CHANGE, 1906-31

Although agricultural production is now a third greater than twenty years ago, crop acreage is only an eighth greater, and quantity of labor employed in agriculture is somewhat less than in 1909. Production per acre has, therefore, increased nearly 20 percent, and production per man nearly 40 percent. Most of this increase has occurred since the World War. The increase in production per acre between 1919 and 1929, two fairly normal years, was about 16 percent, practically none of which is owing to increase in acre-yields of the crops, while the increase in production per man was about 26 percent. It will be noted that agricultural production has just about kept pace with population growth during the past 25 years. (Courtesy, U.S. Bureau of Agricultural Economics.)

S. 44 (*Continued*)

*Descriptive formula:*  $S_{44} = \text{tT}^{-1} : \%(\text{I})_i$

$\%(\text{I})_i = \%(\text{P}, \text{IT}^{-1}, \text{L}^2, \text{PT}_i^{+1}, \text{T}_{ii}^{-1})$   
*Quantic number* = 81 ; 1 ; 2 ; 1

*Legend:*

$S_{44}$ = The situation	P = the U.S. population
records for each of	$\text{IT}^{-1}$ = agricultural production per year
$\text{tT}^{-1}$ = 26 years	$\text{L}^2$ = crop acreage
'  = beginning in 1906	and
$(\text{I})_i$ = 4 indices	$\text{PT}_i^{+1}$ = months of farm labor
$\% $ = in % units	$\text{T}_{ii}^{-1}$ = per year
namely:	

*Comment:*

The situation lends itself to an exploration of tension theory. Let production,  $\text{IT}^{-1}$ , be the desideratum, V, desired by the U.S. population, P.

$$V = \text{IT}^{-1} \quad (\text{a})$$

Let the total intensity of desire be indicated (somewhat indirectly) by the man-days of labor devoted to agricultural production as an index of the national effort to produce:

$$\text{PD} = \text{PT}_i^{+1} \text{T}^{-1} \quad (\text{b})$$

Consider the tension theory equation expanded as a matrix for the series of 26 years:

$$\text{t}(\text{PD} = \text{VE}) \quad (\text{c})$$

The population and the production curves tend to follow each other, i.e., the growth of production correlates with the growth of population:

$$\text{tP} \cdot \text{tV} > 0 \quad (\text{d})$$

$$\therefore \text{tD} \cdot \text{tE} ? > 0 \quad (\text{e})$$

i.e., the average intensity of desire for production must correlate with the national tension in this respect, since, if in (c) P and V correlate highly enough together, D and E must tend to covary also.

The curve for PD (b) is almost level, while P increases through the 26 years, so that D must have decreased compensatingly to produce the constant product PD. When D decreases E must tend to decrease also by (e). The conclusion, based on the assumption (b), is that the national tension towards agricultural production has eased up, less effort in man-days of labor is required as technology increases, surpluses are more apt to occur with the same effort as formerly, abundant agricultural production is less of a problem today to create. The process of detensing,  $-\text{tE}$ , has gone on in respect to this desideratum in this plurel and period.

## IV. NOTES

1. Specifically the formulae are:

$${}_{tT^{-1}} : \begin{cases} \Sigma_v/T = \text{celeration of periodizing, Eq. 70b, Ch. X} & (\text{Eq. 6a, Ch. XI}) \\ \sigma({}_tT) = \text{" " " revarying, Eq. 58, Ch. X} & (\text{Eq. 6b, Ch. XI}) \\ (\partial I)_{\dots} = \text{" " " recovarying, Eq. 65, Ch. X} & (\text{Eq. 6c, Ch. XI}) \\ \sigma I_{t \dots u'} = \text{" " " reranking, Eq. 64c, Ch. X} & (\text{Eq. 6d, Ch. XI}) \end{cases}$$

2. Previous publication of this theory of societal forces may be found in Refs. 12, 16, 18, and discussed in Refs. 6, 7, 44, 55.

3. Since Newton, a physical force has been defined similarly in terms of its net effect, i.e., as the product of a mass,  $M$ , moved a distance,  $L$ , in "time-to-the-minus-two,"  $T^{-2}$ .

$$F' = LMT^{-2} = \text{a physical force} \quad (\text{Eq. 8a, Ch. XI})$$

The unit of physical force is the dyne which is defined as one gram,  $M$ , moved one centimeter,  $L$ , per second per second,  $T^{-2}$ .

For a societal force,  $P$  replaces  $M$  as the population-mass whose inertia the force overcomes. The population becomes that which is moved or changed in some respect. The respect is defined by the index ( $I$ ) which replaces the distance,  $L$ , of the physical force. In  $S$ -notation, mass is a characteristic of people and their environment and is, therefore, a kind of  $I$ , so that the formula for a physical force becomes:

$$F' = T^{-2}I_M L = \text{a physical force in } S\text{-notation} \quad (\text{Eq. 8b, Ch. XI})$$

Thus, we see that physical forces and all their derivatives and compounds, such as "work," "energy," and "power," can be readily expressed as particular non-societal combinations of  $S$ -theory concepts. The indicator in  $S$ -theory can be stretched to include all the variables of any of the sciences, but in order to define a sociological situation, some human beings,  $P$ , must be involved in the situation, explicitly or latently, and the indicators must have some reference to those people.

4. Strictly, this definition of a societal force applies only to phenomena for which objectively observed indices are available. Where such indices are not available the theory may be conceptually applied, but with less precision and assurance. The degree of objectivity is measurable by a reliability correlation coefficient, or a percentage, or some other equivalent index of the degree of agreement between different observers of the same phenomena.

This limitation of the exact application of the theory to situations where precise objective data are available is common to Physics and Sociology alike. Thus, in a collision of two autos, the exact force of the collision cannot be stated until the constituent factors have been measured. If the masses of the two autos and their velocities at the moment of collision are only roughly determined, the force as stated will only be correspondingly approximate. The difference between the physical and the societal forces is that, instead of the easily measured change of distance, the social sciences deal with less easily measured changes of a myriad of characteristics.

5. The three principles as stated for Physics are :

- A. Every body continues in its state of rest or of uniform motion in a straight line, unless it be compelled by impressed force to change that state.
- B. Change of motion is proportional to the impressed force, and takes place in the direction of the straight line in which the force acts.
- C. To every action there is always an equal and contrary reaction.

These principles are currently referred to as the three "laws of motion." The author's suspicion is that this is a misnomer ; they are not laws stating relations between phenomena, such as stated by the law of gravity for example, but simply a generalized operational definition of concepts. The enormous utility of this operational definition has led to venerating it as a "law" of nature's working, rather than as a postulate to help man think about nature's working. As the author is not a physicist, however, this suspicion may be entirely erroneous ; it is recorded to invite criticism from competent persons, as such criticism bears on the philosophy of science and includes our concept of societal forces.

6. Its formula is :

$${}_pF = T^{-2}{}_{\Sigma P}P^p = a \text{ "P-force"} \quad (\text{Eq. 10, Ch. XI})$$

This states that a population force is the acceleration ( $T^{-2}$ ) of a population of some number of persons ( $P$ ), or of plurels ( ${}_{\Sigma P}$ ) of some kind ( $|$ ), raised to some power ( ${}^p$ ). If the exponent is 0, it may specify a sociating force ; if 1, it is a populating force ; if 2, an interacting force. .

7. To derive this, recall that  $P^p = (I^p P_0^p)$  by Eq. 4a, Ch. IV, and that  $P^0 = 1$ , and finally that  $(I)I^p = (I)$ , by Eq. 9b, Ch. III, so that an index in one plurel defined by attribute  $A$  is equivalent to that index for attribute  $A$ , i.e., that index holding under the condition of attribute  $A$ . Attribute  $A$  usually is the name of the plurel, e.g., exports  $(I)$  from Denmark ( $P^p$ ) is equivalent to Danish exports  $(I)$ .

8. Chapin, however, mentions the proposal. (Ref. 6, p. 347.)

9. A formula for causation, based on the definitions above, is here tentatively suggested as an hypothesis :

$${}^t \cdot {}^u ({}_s I_s^s) ? = r = a \text{ formula for the degree of causation} \quad (\text{Eq. 15, Ch. XI})$$

A correlation is asserted here, between an index on one date, or set of dates, and that index on a second date, or set of dates,  ${}^t \cdot {}^u |$ . The conditions are specified by the scripts of the index. Recall that any class script may be the product of any number of implicit attributes which specify the conditions under which the index was observed. The  $r$ , the correlation coefficient, states the probability, the degree to which the earlier index is a cause of the later index. When  $r$  is 1, the probability is perfect, and the earlier index is the sole cause of the later index. The term "cause" here becomes simply a convenient substitute for the cumbersome phrase "antecedent correlate under specified conditions."

This definition of the term "cause" seems to us to have the same referent as the term "mechanism" as defined by Lundberg in Chapter V of Foundations of Sociology. He says: "The word mechanism is used in this book to designate any combination of circumstances, conditions, or movements necessary and

sufficient to produce any observed behavior." The term "necessary" requires a correlation greater than zero between the "circumstances" etc. and the "observed behavior" (the effect) while the term "sufficient" requires a correlation of unity meaning that the "circumstances" etc. are the sole cause of the "observed behavior." We submit that "correlation" covers this case of complete causation and also all degrees of partial causation. It also provides in calculating the appropriate form of correlation coefficient a more operational definition than is conveyed to most persons by the terms "necessary and sufficient." The further element of time sequence in our definition is implicit in his definition in the verb "to produce." His "combination of circumstances, conditions or movements" might be symbolized as an S-situation and not just an index, making Eq. 15 read

$$t \cdot \sigma I = r \quad (\text{Eq. 15a, Ch. XI})$$

This suggestion needs careful exploration and critical revision.

As the application of this formulation (Eq. 15, Ch. XI) to data has not as yet been explored, it may be expected that it will be drastically revised, or rejected as inadequate, upon further research.

10. The generalized standard error formula for these indices building up to a societal force is:

### I. The case of static indices

$$\sigma_I^2 = \sum_1^{t^2} (k_I \sigma P^{-1}) \cdot \cdot \cdot (k_I \sigma P^{-1}) (\cdot \cdot \cdot P \cdot \cdot \cdot \sigma I) = \begin{cases} \text{standard error of a societal force,} \\ \text{acceleration, velocity, momentum,} \\ \text{or change, based on static indices} \\ \text{and a shifting population (Eq. 16)} \end{cases}$$

where: (a)  $I$  can denote (depending on  $k$ )

$$\text{a societal force,} \quad F = PIT^{-2} \quad (\text{Eq. 17a})$$

$$\text{or a societal momentum,} \quad Mm = PIT^{-1} \quad (\text{Eq. 17b})$$

$$\text{or a societal celeration,} \quad = IT^{-2} \quad (\text{Eq. 17c})$$

$$\text{or a societal velocity,} \quad = IT^{-1} \quad (\text{Eq. 17d})$$

$$\text{or a societal change,} \quad \Delta I (= \cdot \cdot \cdot I - \cdot \cdot \cdot I \text{ if } I \text{ is static}) \quad (\text{Eq. 17e})$$

(b) the summation is for all pairs of the  $t$  survey dates,

$$\cdot \cdot \cdot I, \cdot \cdot \cdot \cdot \cdot \cdot \cdot I, \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot I, \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot I, \text{ or more briefly, } \cdot \cdot \cdot, \cdot \cdot \cdot, \cdot \cdot \cdot, \cdot \cdot \cdot \quad (\text{Eq. 18})$$

on each of which  $I$  was measured

$$t = 2 \text{ for a change, velocity, and momentum} \quad (\text{Eq. 18a})$$

$$t = 3 \text{ for celeration and force in 2 consecutive periods} \quad (\text{Eq. 18b})$$

$$t = 4 \text{ for celeration and force in 2 non-consecutive periods} \quad (\text{Eq. 18c})$$

(c)  $\cdot \cdot \cdot \cdot \cdot \cdot \cdot P$  denotes the population of those persons common to both dates (Eq. 19)

(d)  $\cdot \cdot \cdot \cdot \cdot \cdot \cdot I$  denotes the scalar product, in  $\sigma$  units, i.e., the correlation coefficient, of  $I$  on one date with  $I$  on the second date (Eq. 20)

(e)  $k_I$  is a constant in deriving each  $\sigma$ , as it is a ratio of population to time depending on what  $I$  denotes, as follows: For the most complicated

case of a static index and a population whose personnel shifts between survey dates bounding 2 non-consecutive periods:

$${}^{v'}k_F = -{}^{v''}k_F = \frac{.25({}'P + {}''P + {}'''P + {}^vP)}{.5({}^{v'}+{}^{v''}+{}^{v'''}+{}^v)T} \quad (Eq. 21a)$$

e.g., if  $'T = 1900$  A.D.,  $''T = 1910$ ,  $'''T = 1930$ , and  ${}^vT = 1940$ , then  $.5({}^{v'}+{}^{v''}+{}^{v'''}+{}^v)T = (1935 - 1905) = 30$  years

$${}^{v'}k_F = -{}^{v''}k_F = \frac{.25({}'P + {}''P + {}'''P + {}^vP)}{.5({}^{v'}+{}^{v''}+{}^{v'''}+{}^v)T} \quad (Eq. 21b)$$

(When the 2 periods are consecutive, dates  $''$  and  $'''$  become identical,  $'' \equiv '''$ )

$${}^{v'}k_{Mm} = -{}^{v''}k_{Mm} = \frac{.5({}'P + {}''P)}{({}^{v'}+{}^{v''})T} \quad (Eq. 21c)$$

When the population is constant,  $P$  replaces the average  $P$  (which is the numerator above).

For the  $k$  in acceleration, replace the numerator of  $k_F$  by 1; for the  $k$  in velocity, replace the numerator in  $k_{Mm}$  by 1; in change  ${}^{v'}k = 1$  and  ${}^{v''}k = -1$ . (Eq. 22)

The computation of this standard error and its interpretation in special cases can be made clearer by arranging the terms in a square matrix as follows:

Survey dates → ↓		'	''	—	z
	$(k_F\sigma P^{-1})$	$'(k_F\sigma P^{-1})$	$''(k_F\sigma P^{-1})$		$z(k_F\sigma P^{-1})$
'	$'(k_F\sigma P^{-1})$	( $'P$ )	$''P$ ${}^{v'''}I$		$'z_P$ ${}^{v'''}z_I$
''	$''(k_F\sigma P^{-1})$	$''P$ ${}^{v'''}I$	$'''P$		$''z_P$ ${}^{v'''}z_I$
z	$z(k_F\sigma P^{-1})$	$z'_P$ $z'_I$	$z''_P$ $z''_I$		$z_P$

= matrix whose sum is (Eq. 23)

To get  $\sigma_I$ , write the computed values of the  $k\sigma P^{-1}$  term for the appropriate index ( $F, Mm, IT^{-2}, IT^{-1}, I$ ) and for the appropriate date of its determination ( $'', ''', {}^v$ ) as captions of rows and of columns, and write the appropriate computed PI value in each cell, noting that in the main diagonal cells the self-correlations ( $' \cdot 'I = 1.0$ , etc.) are unities reducing that term to  ${}^{v'}P$ , ( $= {}^{v'}P$ ), the population on that survey date. Next in the empty parenthesis in each cell write the product of the three factors: (a) the column  $k\sigma P^{-1}$  term, times (b) the row  $k\sigma P^{-1}$  term, times (c) the cell PI term. The sum of these products in the cell parentheses,  $t^2$  in number, is  $\sigma_I^2$ , and its square root is the standard error of the societal force, momentum, or index used.

Thus, for example, the standard error of a momentum, where the population may shift in personnel between the first and last survey, and where the index is a static one, is:

$$\sigma_{Mm}^2 = (k_{Mm}^2 \sigma^2 P^{-1}) + 2'(k_{Mm}^2 \sigma^2 P^{-1}) + 2'(k_{Mm} \sigma P^{-1})''(k_{Mm} \sigma P^{-1})(''P' \cdot ''I) \quad (\text{Eq. 24})$$

This, for a further, simpler, and familiar example of the standard error of a change (when population is constant so that  $''P \equiv 'P \equiv ''P \equiv P$ ), becomes:

$$\sigma_{I1}^2 = (' \sigma^2 + '' \sigma^2 - 2' \sigma '' \sigma' \cdot ''I) P^{-1} \quad (\text{Eq. 24a})$$

which is the usual sigma of a difference between the 2 means of an index determined on each of 2 dates.

For a non-shifting population, the standard error of a velocity is the standard error of a change divided by the period:

$$\sigma(IT^{-1}) = \sigma({}_t I) {}_t T^{-1} \quad (\text{Eq. 25a})$$

and for a momentum this is multiplied by  $P$ :

$$\sigma(Mm) = \sigma({}_t I)({}_t T^{-1})P \quad (\text{Eq. 25b})$$

as

$${}_t T^{-1}P = k_{Mm} \quad \text{throughout Eq. 24 above} \quad (\text{Eq. 25c})$$

When the index is a percent, since this is a mean and the  $\sigma$  above is the  $\sigma$  of the distribution and not of the mean,  $(\%(100 - \%))^{.5}$  should replace the  $\sigma$  in the right-hand member of the equations above.

The above formulae can be extended by writing appropriate  $k$ 's to give the standard error of a difference between two forces, between two momenta, between two accelerations, between two velocities, or between two changes, so that the statistical significance of any observed difference can be calculated and the probability of its arising by chance as a sampling error may be read off as usual from a table of the normal probability integral.

## II. The case of dynamic indices

All the formulae above are for indices ( $F$ ,  $Mm$ ,  $IT^{-2}$ ,  $IT^{-1}$ ,  ${}_t I$ ) derived from a static index of the characteristic changing. For dynamic indices the formulae are much simpler. The standard error of the dynamic index is the standard error of change, namely:

$$\sigma_{{}_t I} = \sigma_{{}_t I} \text{ of the distribution} / P^{.5} = \text{standard error of a mean change} \quad (\text{Eq. 26})$$

Dividing this by the period gives the standard error of the velocity, as in Eq. 25a, and multiplying this by the population gives the standard error of the momentum, as in Eq. 25b. The standard error of the acceleration (for a constant population) is that of a difference of two mean dynamic indices (paralleling Eq. 9a) divided by the period between middates of the two velocity-periods:

$$\sigma(IT^{-2}) = (P^{-1}({}_t \sigma^2 + {}_{t'} \sigma^2 - 2{}_t \sigma {}_{t'} \sigma' \cdot {}_t I))^{.5} \cdot {}_t T^{-1} \quad (\text{Eq. 27})$$

The standard error of the force derived from dynamic indices is this standard error of the acceleration multiplied by the constant  $P$ :

$\sigma_F = \sigma(IT^{-2})P$  standard error of a societal force based on dynamic indices and a constant population (Eq. 28)

When the population shifts between the two periods, the formulae for the standard error of the acceleration and of a force based on mean dynamic indices of change is more complicated. To derive it, write a force as the time rate of change of the two momenta:

$$F = ({}_{v'}(PIT^{-1}) - {}_{v''}(PIT^{-1})) \cdot {}_{.5(v+''-''-')}T^{-1} \quad (\text{Eq. 29})$$

Let  ${}_{,k} = -{}_{,T^{-1}} \cdot {}_{.5(v+''-''-')}T^{-1}$ ;  ${}_{,,k} = {}_{,,T^{-1}} \cdot {}_{.5(v+''-''-')}T^{-1}$  (Eq. 30a and b)

Then  $F = {}_{,k}PI + {}_{,,k}PI \quad (\text{Eq. 31})$

The usual process of deriving a standard error formula is to differentiate both sides of Eq. 31, treating  $k$  and  $P$  as constants, square, sum, and divide. These steps exactly parallel Eqs. 7 to 13 in Ref. 16 (but note misprint there in the date script in Eqs. 3 and 5, where  $(4 - 3 - 2 + 1)$  should read  $(4 + 3 - 2 - 1)$  as in  $({}_{v+''-''-'})$  here). The result is:

$$\sigma_F^2 = {}_{,k}\sigma^2 + {}_{,,k}\sigma^2 + 2({}_{,k}\sigma)({}_{,,k}\sigma) \cdot {}_{,,P} \cdot {}_{,,I}, \text{ standard error of a societal force,} \\ \text{based on dynamic mean indices} \\ \text{and a shifting population} \\ (\text{Eq. 32})$$

${}_{,,P}$  denotes that part of the population which is identical in both periods.

In all the standard error formulae above, it is assumed that the conditions of random sampling hold. The interpretation of these standard errors is, as usual, in probability terms where, if an index is more than three times its standard error, it is conventionally considered reliable, as it would occur by chance once in a thousand times; if the index is four times its sigma, the chance element is thirty-two in a million; if five times, three in ten million; if six times, one in a billion.

11. Before concluding the exploration of this chapter into the compounds of acceleration, it may be worth-while to consider carrying the compounding of time, indicators, people, and space still further, as Physics does, to develop useful concepts of societal "work," "energy," "power," "pressure," etc. As yet no quantitatively recorded situations in the literature of the social sciences have been found which require such compounds to describe them. The policy throughout this volume is to contribute to the building of a systematic science of Sociology by induction from observed facts, not from armchair speculation. The 1500 graphs are the factual base thus far. Every concept, every formula proposed, represents some facts; and every fact (i.e., every graph) is described and classified by these concepts. No concepts or formulae have been invented for which no corresponding facts were found; permutations and combinations of the sixteen basic concepts will be developed as far as needed to cover all the facts. "Force" and "control" are factually based inductions as evidenced by the 41 situations appended to this chapter. With further research, the compounds of "work," "energy," "power," "pressure," etc., of Physics, may prove useful and factually based. But since no situations corresponding to their

formulae have been found, fuller discussion of such societal concepts is refrained from, and attention is merely called to their formulae as a field for further research.

A force accomplishing a societal change may be called "*societal energy*" and may be defined as the product of the force and the change:

$$FI = (T^{-2}IP)I = En = \text{Energy} \quad (\text{Eq. 34, Ch. XI})$$

This is the product of the population and the square of its velocity of change. If this changing is thought of as finished, i.e., in the past tense, it may be called societal *work* done. If it is a current change, it may be called kinetic energy; while if it is thought of as possible in the future, it may be called potential energy. (Note that, since the change correlates perfectly with itself, the arithmetic product above is identical with the scalar product, and the quantic number is 8;2;0;1.)

*Power* is the time rate of doing work:

$$En T^{-1} = T^{-3}I^{+2}P^{+1} = Pw = \text{Power} \quad (\text{Eq. 35, Ch. XI})$$

$$|^s = 7;2;0;1$$



## *PART VI*

### THE SYNTHESIS OF SECTORS, S

*studying situations defined by  $S = \textcircled{(T;I;L;P)}$*



## Chapter XII

### APPRAISAL OF S-THEORY, S.

#### I. SCIENCE AS UNDERSTANDING, PREDICTING, AND CONTROLLING PHENOMENA

The detailed exposition of S-theory has been completed in the preceding ten chapters. The variations, permutations, and combinations of its sixteen chief concepts (the four sectors, the four scripts, and the eight essential operators— $T I L P \begin{smallmatrix} \S \\ \S \\ \S \end{smallmatrix} : :: \bullet ' \pm \times / ^1$ ) have been systematically explored, and the degree to which these combinations describe societal situations and provide reliable and precise sociological concepts has been exposed. In conclusion an appraisal of S-theory as a contribution to the science of Sociology will be offered. This appraisal is, of course, tentative, both because it is made by the author who is not an impartial judge of his own creation and because it is made too early. For a mature appraisal of its contribution, its use or non-use by sociologists, in whole or in parts, over a long period is required.

To appraise this system, the function of science must be recalled. The postulate here is that the function of science is to aid man in understanding, predicting, and controlling phenomena. Understanding, predicting, and controlling are components of the process of adjusting. Science is the most effective cultural technic for adjusting in our physical and social environment that man has yet developed, and whatever is believed by men to be a satisfactory adjustment is for those men at that time the summum bonum of life. A "satisfactory adjustment" is a statement, in psycho-biological language, which can include formulations in other languages, of the purpose of human life. Thus, in the theological view that "the chief end of man is to glorify God," the "satisfactory adjustment" is conceived as an adjustment towards whatever represents God in that believer's psychic and physical environment. A satisfactory adjustment may be conceived in philosophic terms as the quest for truth, beauty, and goodness.

In the utilitarian view that the summum bonum is "the greatest happiness of the greatest number," a more perfect democracy would be a more satisfactory adjustment. Fascism, communism, and other ideological systems of supreme values, can be translated as satisfactory adjustments to their adherents. Men will doubtless continue to differ as to the purpose of life, the summum bonum, and a consensus from human experience may or may not be in process, but there is almost world-wide agreement that science is one outstanding means towards whatever ends humanity desires. It is one means among others for achieving whatever man desires, including unethical as well as ethical desires. Science is a tool, not an end in itself. Thus, scientific theories such as S-theory are seen as attempts to help forge better tools for humanity to use.

The excellence of the tool, science, is testable by the excellence of the understanding, prediction, and control of phenomena which it gives to man. The contribution of S-theory to each of these three functions will be discussed in turn.

## *II. CONTRIBUTION OF S-THEORY TOWARDS UNDERSTANDING PHENOMENA*

We take the view that understanding phenomena consists in the conceptual organizing of sense perceptions, or in behavioristic language, the hierarchical organizing of responses to stimuli. If some phenomenon seems consistent with our current conceptual organization of sense perceptions we say, "We understand it." If the phenomenon is inconsistent with such a conceptual system, we respond to it without understanding it. Our knowledge derives from sense perceptions. Seeing relations between these organizes them into some sort of a system. Among the simplest relations are those of similarity which yield concepts, i.e., generalized percepts. "A man" is a concept, generalized from many perceptions of particular men. "Time" is a concept generalized from particular experiences of intervals between events and from amounts of activity which may initially be the internal physiological activity of heart, lungs, and other organs, but becomes objective by means of clocks and other instrumental readings. An "indicator" is an explicit generalization of every record of a qualitative or a quantitative characteristic, i.e., item of knowledge, other

than time, space, and population. A first step in understanding phenomena is to construct concepts generalizing our percepts of phenomena of one kind, or class, and increasingly to construct concepts with operational definitions to improve their standardization.

### *A. The Contribution of S-concepts in Understanding Society*

The concepts defined in this volume and represented by the sixteen basic symbols ( $T I L P \frac{S}{S} : :: \bullet' \pm \times /$ ) and their derivatives are claimed to be: (a) objective, (b) precise, (c) parsimonious, and yet (d) comprehensive.

## 1. OBJECTIVITY OF S-CONCEPTS

Their objectivity has been experimentally measured by the 93%, 97%, and 99% of agreement between independent analysts using them in controlled experiments as described in Chapter II. Has any other system of sociological concepts been subjected to similar scientific testing? There remains, however, the challenge to develop concepts whose objectivity approaches 100% still more closely. Hereafter any system of concepts which is offered without experimental testing of their objectivity should be regarded by sociologists as scientifically deficient. Psychologists have come to regard the publication of a new scale without determination of its reliability as an incomplete job, falling short of scientific standards. Chemists must know the degree of purity of their reagents, physicists must know the degree of precision of their instruments. Sociologists, if they wish to be scientists and not literateurs, should no longer naively ignore the inaccuracy of their conceptual tools when technics for determining such accuracy have been developed and demonstrated. In experimentally demonstrating the objectivity of S-theory, it is claimed that the conceptual tools of sociology have been improved, thus contributing to the understanding of the phenomena which these concepts summarize.

## 2. PRECISION OF S-CONCEPTS

It is submitted, secondly, that the concepts of S-theory have a high degree of precision in their definition. They are defined by

equations of mathematically combined symbols. These symbols for the most part represent quantitative or quantifiable phenomena. Even the remainder that seem intrinsically qualitative are precisely dealt with in the operational formulae defining them. The list of the five hundred odd equations of this volume constitutes a list of the precisely defined concepts of S-theory.

In contributing increased precision to societal data, the use of algebraic symbols is important. They strip the words naming current concepts of their subjective and emotional connotations, leaving the essential agreed-upon denotation of the concept. The mathematical and logical rules for manipulating algebraic symbols are more precise, and therefore, distinguish between truth and falsehood more exactly than the grammatical rules for manipulating words in sentences.<sup>2</sup>

### 3. PARSIMONY OF S-CONCEPTS

Thirdly, the concepts here offered are believed to be a parsimonious set. No compounded concepts have been introduced unless situations requiring them were discoverable. The compounded concepts are special cases or combinations of sixteen basic concepts—time, space, people, characteristics, the exponent, class, class-interval, case, aggregation, cross-classification, correlation, identification, addition, subtraction, multiplication and division  $\mathfrak{s}(T;I;L;P)\mathfrak{s}$  ;  $= (: :: \bullet' \pm \times /)$ . The four arithmetic operations of addition, subtraction, multiplication, and division are common to any scientific system and not peculiar to S-theory. The other operational symbols,  $\bar{\phantom{x}}$ ,  $\_$  ? (see Glossary of Symbols of S-theory, Appendix I) are convenient but dispensable. With the possible addition of the subindicator of desiderata, the sixteen concepts above are the essential base of our system.<sup>3</sup>

### 4. COMPREHENSIVENESS OF S-CONCEPTS

The fourth claim for the concepts of S-theory is that they are a comprehensive set. As evidence for this claim two comparisons may be made—with Eubank's list of 332 current concepts and with the list of 61 recommended as basic by the Committee on the Introductory Course in Sociology (Ref. 10). Eubank's "Selected Catalog of Terms Used as Concepts of Sociology as

Found in the Literature of Sociological Theory" in his book on the *Concepts of Sociology* (Ref. 25, pp. 39-43) has been roughly translated into S-theory concepts. Out of the 332 terms only 44 (13%) cannot be expressed in S-terms—terms such as "conation," "consciousness," "discussion," "mind," "objectification," "positivism," "reality," "sociology," "subjectification," "thought," "will," etc. 78 terms are precisely redefined by S-equations (23%). The remaining 210 terms can be expressed in S-theory concepts to the extent that indicators, better than mere names, can be found to represent their meaning (i.e., as far as indicators such as a scale, list of acts, or other objective operational ways can be devised to define such concepts as "attitudes," "achievement," "attraction," "compliance," "estrangement," "intimacy," "mores," "prejudice," "repulsion," "status," "unrest," at least as used in specified situations or under stated conditions). In short S-theory can express 87% of the list of terms with an adequacy which depends not on S-theory but on the adequacy of existing objective technics of determining that which those terms denote.

A second set of evidence of the comprehensiveness of the concepts of S-theory is a comparison with the list of 61 terms recommended by a committee of the American Sociological Society as standard content for the introductory college course in Sociology. The list is reproduced here with an indication of their equivalents in S-theory:

accommodation	Eq. 30, Ch. X	caste	Eq. 4, Ch. IV; Eq. 100, Ch. X
adaptation	Eq. 93c, Ch. X	collective be-	$P_p : (IT^{-1})^{\frac{1}{2}}$
adjustment	Eq. 93c, Ch. X	havior	
amalgamation	Eqs. 6a and 26, Ch. X	community	Eq. 16, Ch. VII
assimilation	Eq. 26, Ch. X, also Eq. 96, Ch. X	communication	Eq. 12, Ch. X *
association	Eq. 6a, Ch. X, also dissociation Eq. 66, Ch. X; Eq. 8, Ch. IV	competition	Eq. 47a, Ch. X
		conditioned response	Eq. 25, Ch. X *
		conflict	Eq. 10, Ch. X
attitude	Eq. 39, Ch. X; see S. 12, Ch. II; S. 11, Ch. III; S. 1, 2, 8, 14, Ch. V	contact	Eqs. 12, 13, 14, Ch. VII
		co-operation	Eq. 32, Ch. X
		crowd	Eq. 10b, Ch. X
		cultural change	Eq. 27, Ch. X
behavior	Eq. 27b, Ch. X	cultural lag	Eq. 63, Ch. X
pattern		culture	Eq. 13, Ch. III

culture area	See S. 1, Ch. VIII	leadership	Eqs. 8, 9, Ch. VII
culture complex	Eq. 13, Ch. III, also Eq. 15, Ch. IX	mores	Eqs. 26, 34, Ch. IV; Eq. 33, Ch. X
culture pattern	Eq. 13, Ch. III, also Eq. 15, Ch. IX	personality	Eq. 27, Ch. X *; Eq. 13, Ch. III
culture trait	Eq. 13, Ch. III, also Eq. 15, Ch. IX	primary group	Eq. 4, Ch. VII *
custom	Eq. 15, Ch. IX	progress	Eq. 34, Ch. X; Eq. 27, Ch. V
diffusion	Eq. 78, Ch. X ("mo- mentum"), also Eq. 67cI	race	Eq. 3, Ch. IV *; II C, Ch. IV
disorganization	Eq. 102, Ch. X; Eq. 26, Ch. IV	secondary group	Eq. 4, Ch. VII *
ethnocentrism	P <sub>p</sub> : S * = data cor- responding to ra- cial plurels	society	Eq. 4, Ch. VII
folkways	Eq. 15, Ch. IX	social class	Eq. 3, Ch. IV
geographic de- terminism	See S. 10, 12, 16, 17, 18, 19, 21, 22, Ch. VIII	social control	Eq. 16, Ch. XI
geographic or physical or natural en- vironment	See formulae for S. 6, 10, 12, 13, 18, 21, 22, Ch. VIII	social distance	Eq. 4, Ch. VII *; see S. 12, Ch. II
group	Eq. 4, Ch. VII; Eq. 10c, Ch. VIII; Eq. 1d, Ch. IV	social or psy- chological or cultural en- viron	Eq. 62, Ch. X *; also see S. 1, 2, 4, 17, 21, 22, Ch. VIII
human nature	Eq. 13, Ch. III; Eq. 27, Ch. X *	social evolution	Eq. 97, Ch. X
imitation	Eqs. 80, 81, 82, 83, 84, 85, 86, Ch. X	social heritage	Eq. 14, Ch. VII *; Eq. 58b, Ch. X; Eq. 87, Ch. X
instinct	Eq. 37, Ch. X *	social interac- tion	Eq. 12, Ch. X; Eqs. 1, Ch. VII
institution	Eq. 13, Ch. III; see S. 4, Ch. III	social organiza- tion	Eq. 102, Ch. X; Eqs. 3, 4, Ch. IV
invention	Eq. 25a, Ch. X	social process	Eq. 2, Ch. X
isolation	Eqs. 10, 11, Ch. VII	socialization	Eq. 93, Ch. X
		status	Eq. 26, Ch. V
		stratification	Eq. 100, Ch. X
		values	Eqs. 27, 33, Ch. X; Eqs. 25, 34, 49, Ch. V

The 15 starred terms, or about 25% of this list, can be expressed, but not advantageously, in S-terms by suitable defining of indicators and scripts, while the remaining 75% have clear-cut formulae redefining them more precisely than hitherto. Three quarters of the list recommended as a standard set of sociological concepts have their definitions or descriptions <sup>4</sup> improved in objectivity, precision, and measurability by the S-theory!

In addition, the following 120 terms of limited but increasing currency in the sociological literature are defined or have their form specified by S-equations:

abnormality	Eq. 26, Ch. V	economic process	Eq. 23a, Ch. X
acceleration of change	Eqs. 4, 6, Ch. IX	elements	Eqs. 13-19, Ch. VI
adequacy of sampling	Eq. 21a, Ch. III	energy	Eq. 17, Ch. XI
amelioration	Eq. 29, Ch. V; Eq. 66b, Ch. X	equalitarian	Eq. 16, Ch. VII
attracting	Eq. 9, Ch. X	equalizing	Eq. 45b, Ch. X; Eq. 28, Ch. V
attribute	Eq. 18, Ch. II; Eqs. 3-8, Ch. III	error	Eqs. 45-48, Ch. VI
behavior	Eq. 32, Ch. V; Eq. 29, Ch. X	evolving	Eq. 97, Ch. X
cardinal cases	Eq. 16, Ch. III	exploiting	Eq. 96, Ch. X
	Eqs. 33-37, Ch. II; Eq. 12, Ch. IV	exponent	Eq. 1, Ch. II
causation	Eq. 15, Ch. XI	force, societal	Eq. 7, Ch. XI
change	Eq. 5a, Ch. X	heredity	Eq. 58b, Ch. X
class	Eqs. 23-26, Ch. II	hierarchy	Eqs. 40, 41, Ch. II; Eq. 13, Ch. III
class-interval	Eq. 27, Ch. II	identification	Eqs. 48, 49, Ch. II
classification	Eq. 40, Ch. II; Eq. 3, Ch. IV	imitation	Eqs. 80-86, Ch. X
coercing	Eq. 40a, Ch. X	index	Eqs. 10, 12, Ch. III
commercializing	Eq. 98, Ch. X	indicators	Eqs. 3, 9, 12, Ch. III
communality	Eq. 45, Ch. VI	individualizing	Eq. 92, Ch. X
components	II A, B, Ch. VI; Eqs. 20-22, Ch. VI	integrating	Eq. 102, Ch. X
contingency	Eqs. 49-51 of S. 11, Ch. VI	interaction	Eq. 12, Ch. X
correlation	Eqs. 1-5, Ch. VI; S. 1-24, Ch. VI	interrelation	Eq. 14, Ch. VII
cosine	Eq. 30, Ch. III	kinship terms	Eq. 14, Ch. IV
cross-classification	Eq. 45, Ch. II	larithmics	Eq. 9, Ch. X *
cycle	Eq. 3, Ch. XI	lead and lag	Eq. 62, Ch. X
dates	Eqs. 1b, 2f, 3d, Ch. IX	leadership	Eq. 89, Ch. VII
density	Eq. 7, Ch. VIII; Eq. 3, Ch. II	liberating	Eq. 40b, Ch. X
desiderata	Eq. 51, Ch. V	limits	Eq. 20, Ch. III; Eq. 35, Ch. II
desire	Eq. 35, Ch. V	maps	Eq. 3, Ch. VIII
determination, coefficient	Eq. 34, Ch. II	matrix	Eq. 19, Ch. III
deviation	Eq. 24d, Ch. III	mean	Eq. 8, Ch. V; Eq. 12, Ch. V
differentiation	Eq. 14, Ch. III	migrating	Eq. 22a, Ch. X
dimension, societal	Eq. 52, Ch. II; Eqs. 35, 37, Ch. III; Eq. 18, Ch. IV	mobility, gross and net	Eqs. 13 and 14, Ch. X
dispersing	Eq. 45a, Ch. X	moments (statistical)	Eqs. 7-11, Ch. V
distribution	Eqs. 1-5, Ch. V	momentum	Eq. 78, Ch. X
durating	Eq. 71, Ch. X	monopoly	Eqs. 19 and 20, Ch. V; Eq. 49, Ch. X
dynamic	Eq. 3, Ch. IX	normal probability	Eq. 15, Ch. V
		normality	Eq. 10, Ch. VI; Eqs. 14 and 26, Ch. V
		observation error	Eqs. 21 and 22, Ch. III
		operators	Eq. 50, Ch. II
		opposing	Eq. 95, Ch. X

ordinal	Eq. 16, Ch. III	situation	Eq. 50, Ch. II
origin	Eq. 20e, Ch. III	skewing	Eq. 10, Ch. V; Eq. 66, Ch. X
percent	Eq. 17, Ch. III	social distance	Eq. 6, Ch. VII
periods	Eq. 29, Ch. II	margin	
persecuting	Eq. 99, Ch. X	social problem	Eq. 26, Ch. V
persons	Eq. 37, Ch. II	stability	Eq. 58, Ch. X
plurels	Eq. 4b, Ch. IV	standard deviation	Eqs. 9 and 18, Ch. V
populating	Eq. 9, Ch. X	standard error	Eq. 56, Ch. VI; Eqs. 13, 16, 18, Ch. V
population pyramid	Eq. 12, Ch. IX	static	Eq. 3, Ch. IX
power	Eq. 18, Ch. XI	subclassification	Eqs. 39-43, Ch. II
probability	Eqs. 7-9, Ch. VI	subnormal	Eq. 26b, Ch. V
profiting	Eq. 24, Ch. X	tension	Eq. 35, Ch. X; Eqs. 34, 36, Ch. V
proselytizing	Eq. 92a, Ch. X	time series	Eq. 43, Ch. II
quantic	Eq. 21, Ch. II	tolerating	Eq. 92, Ch. X
range	Eq. 30, Ch. V; Eq. 92, Ch. VI	trend	Eq. 2, Ch. XI
ranks	Eq. 18, Ch. III	types	II B 5, and Eq. 100, Ch. VI
reliability	Eq. 64, Ch. X; Eqs. 21 and 22, Ch. III	unifying	Eq. 90, Ch. X
remapping	Eq. 69, Ch. X	validity	Eq. 23, Ch. III
representative sampling	Eq. 23d, Ch. III	valuating	Eq. 30, Ch. X
revolution	Eq. 5, Ch. XI	variance	Eq. 9, Ch. V
rural-urban	Eq. 9, Ch. VIII	varying	Eq. 21c, Ch. V; Eq. 94, Ch. X
sampling error	Eq. 14, Ch. IX; Eq. 21, Ch. III; Eq. 16, Ch. V; Eqs. 1-17, Ch. XI, Appendix III	vectors	Eqs. 26-31, Ch. III
sequence	Eq. 1, Ch. X	velocity of change	Eq. 5d, Ch. X; Eqs. 3 and 5, Ch. IX

The 46 terms from the Committee's list and these 120 additional terms make up a body of some 166 compounded terms. Another four hundred more technical terms<sup>5</sup> of less frequent occurrence are also defined or specified by the S-formula in the equations of this volume. This list of more than five hundred and fifty terms in all, evidences the comprehensiveness of the twelve constituent distinctive concepts of S-theory together with their combinations.

Comprehensiveness, however, is but one of the four properties of concepts, along with objectivity, precision, and parsimony, all of which, according to our argument, tend to increase human understanding of phenomena.

### *B. The Contribution of the S-system in Understanding Society*

Understanding phenomena consists further, in the view as stated above, of the *organizing* of concepts derived from percepts, i.e., in arranging knowledge into a consistent system.

#### 1. THE QUANTIC SYSTEM

What is the contribution of S-theory towards organizing sociological concepts into a consistent system? Here is where the theory makes perhaps its most significant contribution towards building up a science of Sociology, for S-theory provides in a single mathematical formula a system of which attributes and variables, qualitative and quantitative occurrences, probabilities, distributions, correlations and stimulus-response interrelations, time trends of change and societal forces, in short, time, space, people, and their characteristics are all articulated parts. All the concepts of prediction and probability and of sociological concepts of content, listed above, are integrated in a consistent, closely unified system, which is specified by the S-theory equation,  $S = {}_s(T;I;L;P)_s$ . Every one of the hundreds of concepts defined by a formula in this volume can be shown to be a special case, a particular permutation and combination of that general formula. Is there any other system of Sociology that can claim as much? Is there any sociological treatise or textbook that can show that its chapters and the topics treated are derivatives of one formula, or are such logically necessary parts of one system that by recombining their basic concepts systematically, or even in random ways, these new combinations yield the topics treated?

Towards the aim of science to bring order out of chaos, to systematize our universe, the claim is made here, for sociologists to judge, that S-theory, whatever else its inadequacies may be, does develop such a *system* for the sociological segment of knowledge better than any other sociological system offered to date.

#### 2. THE PARALLELING GEOMETRIC SYSTEM

The function of S-theory in systematizing societal phenomena is particularly evident in connection with correlation in its geometric interpretation. With the correlation coefficient between any two indices determining the angle between their vectors,

and with all quantitative phenomena representable by a sheaf of such vectors in  $n$ -dimensional space, a geometric model can be made for all quantitative societal data (see S. 35, Ch. II). This model visualizes the dimensions of society. This model, extended as in topology where angles and lines are not yet measured, can include qualitative phenomena as points in  $n$ -space, but points between which distances and directions are mostly undetermined as yet. The internal consistency of the system which is S-theory can then be tested by the rigorous reasoning of geometry.

The geometric system may be summarized as a hierarchy of entities in  $n$ -dimensional space. As described in the diagram at the end of the text of Chapter II this hierarchy may be presented in either geometric or sociological terms. Geometrically, space is subdivisible into observed subspaces called "situations," which are subdivisible into four sectors, defined by vectors and their normals, and these are further subdivisible into sects and points. Sociologically, society is observed in samples called situations, which are analyzable into indices of four general kinds, which can be operationally related to each other and which are subdivisible into quantitative units or points. Symbolically in S-notation, any S is a combination of four kinds of indices ( $I'$ ) with a specified number of each kind,  $|_s$ , and specified exponents,  $|^s$ , with specified class-intervals,  $|_s$ , and points,  $|^s$ . In this hierarchy the situation is the gross sociological unit, the index is the simplest qualitatively distinguishable constituent unit, and the class-interval is the quantitative unit within the index. All these are specifiable entities. The term "dimension" is a convenient name for the indices including the denotation of both the class script and the exponent. Popularly it connotes a measurable or quantitative entity; algebraically it connotes the exponent,  $|^s$ ; geometrically it connotes vectors,  $|_s$ . The dimension is perhaps the fundamental concept in S-theory, though in practice it is more conveniently analyzed into observed indices and their relations separately. Those relations are as denoted by the exponent which expresses the degree of similarity of indices (as  $I^2$ ), or puts one into terms of another (as  $T^{-1-2}$ ) or combines them in some form of multiplicative product (as  $P^2$ ,  $L^2$ ,  $I^2$ ). The number of dimensions in a recorded societal situation is the total number of qualitatively distinguished entities in it. The concept of a dimension

includes a kind of something, an amount of it (including unity and zero), and relations to other dimensions. Visually, dimensions are lines (including the special case of merely two points) of certain lengths and with certain angles between them.

The systematization which S-theory provides can also be visualized in S. 33, Ch. II, the quantic solid. Here is diagrammed the quantic classification of all quantitatively recorded societal situations. According as the exponents vary in each of the four sectors, representing the operational thoroughness of observing time, space, people, and their characteristics, every situation (specified as a particular observed unit-combination of indices from these sectors with their exponents) is classified by its quantic number into one of some 81 most frequently occurring cells in this quantic scheme. The reliability of this classifying of situations has been experimentally demonstrated to be 97% of agreement between two independent analysts in one experiment, 99% in another experiment, and 98.5% on repetition after a month by one analyst.

This clear-cut classifying of societal situations, of observed portions of the life of human society, by their quantic numbers is believed to be a unique contribution of S-theory among systematizing theories in the social sciences. Its full significance cannot be appraised so early. Like Mendeleev's classification of the atomic elements in Chemistry into the periodic table, the attendant properties of each class and family of classes may cumulatively be discovered and developed with the research of decades, so that the utility of the classification may become far greater than its bare neatness gives promise of at first. The classification does not follow conventional thoughtways and will consequently seem to lack "meaning" to many sociologists and social scientists. This is as inevitable as Mendeleev's classification basing itself on properties of weight and valence which proved to be more fundamental than the previously familiar color, smell, medicinal uses, etc. of his chemical elements. But in proportion to the objectivity, the precision, the parsimony, and comprehensiveness of the concepts of the classification, it may be expected to be used increasingly by the on-coming generation of sociologists, for whom its "meaning" will be the associations acquired through using it in various ways.

### 3. HYPOTHESES IN THE SYSTEM

It is, of course, to be expected that this system will not continue as formulated in this volume. It should grow and develop and perhaps change almost unrecognizably if it is a good theory. Or again, parts of it may prove useful while the rest may be forgotten. Thus, in the system of hypotheses that make up S-theory some few hypotheses may prove to have more enduring utility than others, such as, possibly:

1. the hypothesis of sectors—Ch. II, IB
2. the hypothesis of societal vectors—Ch. III, VI, and Ch. VI
3. the hypothesis of the quantic classification—Ch. II, IC1
4. the hypothesis of attributes expressible as  $I^\circ$ —Ch. III, II A, B
5. the hypothesis of precision of measurement—Ch. III, IVB
6. the hypothesis of a natural range of  $12.5\sigma$ —Ch. V, IID
7. the tension hypothesis in the tension theory—Ch. V, IIE and Ch. X, IIB2
8. the hypothesis of epsilon elements—Ch. V, IF1, Ch. VI, IIC
9. the hypothesis of a group and of a community—Ch. VII, I and IIID2
10. the hypotheses of societal processes—Ch. X
11. the hypothesis of societal force—Ch. XI, II
12. the hypothesis of societal control—Ch. XI, III.

It is possible that the notation: (a) of the indicator with a zero exponent bringing qualitative data within mathematical rules; (b) of the descripts; and (c) of the colon denoting orderly aggregation, such as in any classification, may prove to have outstanding utility.

These hypotheses develop the concepts of S-theory and relate them together systematically, towards increasing our understanding of societal phenomena.

#### *C. Limitations of S-theory in Understanding Society*

Among the many limitations of S-theory which the student may note, the following are apparent to the author and are listed here in summary:

1. In its *content* the theory is limited (a) to records of societal

phenomena. (b) These, for the most part, must be in quantitative or at least itemized form, since the extent to which the S-theory covers qualitative non-itemized data is not fully explored as yet. (c) The boundary of such data includes all the social sciences and does not mark off as peculiar a field to Sociology, as many sociologists would like to see delimited. (d) The theory has not yet been extended (by developing appropriate indicators) into each of the social sciences in a systematic way, though this seems feasible.

2. In its *form*, the S-formulae are (a) descriptive more than calculative. They describe the form of the societal situations analyzed and usually do not enable the calculation of unknown variables in the situation. The S-formulae serve (b) to classify societal phenomena more than to reveal their functioning.

3. The *reliability* of the theory has only been tested by a few experiments and these show unreliability of 1 to 3% for the quantic formula and of 7% for the full descriptive S-formula.

4. The *validity* of the theory has only been tested by three analysts on a sample of 1500 situations, all of which it fitted without exception. On a larger sample of societal situations an exception, not analyzable into an S-formula, might be discovered.

5. In the *presentation* of S-theory in this volume there are many deficiencies, among which are:

- a. Many formulae lack rigorous derivations (e.g.,  $P^2$  denoting interrelation of parties); and the standard errors of most newly proposed indices have not yet been derived.
- b. Many of the hypotheses proposed are so briefly discussed as to be inadequately explored here.
- c. Overclaims have doubtless been made or implied partly because the limiting qualifications of many formulae and hypotheses here have not yet been determined. (E.g., perhaps the derivation of all economic data from the interrelation matrix; and perhaps the claim that all qualitative itemized characteristics can be represented by  $I^0$ .)
- d. The algebraic symbolism of S-theory is too detailed for many readers, or, alternatively, their interest and skill in reading these formulae have not been successfully developed in the early exposition of the symbols.

- e. A logical rather than a psychological plan was chosen for the chapters and sections of this book. The S-formula was systematically explored with consequent dryness probably to the reader rather than picking out its interesting applications and ignoring the rest. The aim was to contribute to the understanding of the professional social scientist rather than to popular understanding at first.

### III. CONTRIBUTION OF S-THEORY TOWARDS PREDICTING PHENOMENA

#### A. Classes of Predictions

Prediction is foretelling future situations.<sup>6</sup> \* Predictions may be classified either on the basis of the time factor as static or dynamic, or on the basis of the participation of the predictor as participant or observant predictions.

#### 1. PARTICIPANT PREDICTION

Observant prediction is the rule in sciences other than the social sciences, but participant prediction becomes important in the social sciences, for here the prediction may be an intention, a statement of purpose, which the party predicting then works to make come true, and which without his effort might not be realized. The predictor's knowledge and desire and consequent behavior may be a cause of the predicted effect. Participant prediction is internal, purposeful; while observant prediction is external and may, or may not, involve purpose, depending on whether any people's desires are causes of the predicted situation. The former is predicting by intent, the latter by portent. Both are based on probability. Both can be behavioristically observed without mystic telic subjectivity. Thus, an insurance company may determine the probabilities of longevity of persons of given ages and amounts of overweight. If those persons learn that their overweight tends to reduce their chances of a desired long and healthy life, they may tend to correct their fatness and so alter the prediction of their longevity. This simply means that new actuarial data must be secured determining the new probabilities for people who know how to, and desire to, modify the former prediction which held only on the condition of their ignorance or apathy.

\* For Eq. 1, Ch. XII, see notes at end of chapter.

The probabilities for participant prediction may be different from the probabilities for observant prediction, but both can be actuarially determined. Sometimes the participant prediction may be determinable with greater precision. Thus the accounts of a person, a family, a company, or a nation, are more accurately predictable if that party has adopted a budget and thus participates in creating the predicted outcome. The modern growth of social planning by nations, communities, or other groups is a major trend towards more participant prediction.<sup>7</sup> Any publicly announced goal, if genuinely desired, may be a stimulus to effort, roughly in proportion to the distance (within limits) from its achievement. Thus, the adoption of a goal may be a major cause of its fulfillment, and greatly increase the probability of that prediction. The fatalistic notion that prediction of human affairs is somehow impossible because "purpose" enters in, because the "personal equation is a factor," because "an unpredictable telic element is involved," should be replaced by the working hypothesis that human purposes are objectively observable by suitable indicators of speech and action, and that actuarial data can be collected, yielding progressively higher probabilities for such participant predictions. The issue is at bottom one of adequacy of factoring a situation as discussed in the preceding chapter in connection with the Gestalt theory. In participant predictions failure to observe indicators of the purposes (i.e., intensities of attitude) of any participants in the predicted situation is simply inadequate factoring of that situation, and the degree of inadequacy is measured by the difference between the probability of the situation (as inadequately observed) coming true and perfect probability. Such deficits from perfect probability in its various forms from a simple ratio to joint probability, to normal probability, to the alienation coefficient, etc., as sketched in Chapter VI on Correlation, have been systematically presented in S-theory terms as compounds, defined by operational formulae, of the sixteen basic <sup>8</sup> S-theory concepts.

## 2. STATIC PREDICTION

Predictions may also be classified on the same basis as the quantic classification and reviewed by following the chapters of this book. This starts with predictions from static data, first

from indicators, then from plurels, then from distributions, correlations, interrelations, and densities, followed by dynamic predictions from durations, velocities, and accelerations.

As types of prediction from indicators alone, consider Chapin's tabulation of the levels of symbolic substitution, S. 1, Ch. XII. Each ascending level from the bottom "0 level" up to the "6th level" increases in generality. Under the conditions where any one of those observations were made there is a probability that similar events would recur, i.e., seeing an unemployed person again the next day, (0 level, 2nd column), etc. The probability increases with the symbolic generality until at the top level the equation asserts perfect probability, i.e., that for observed values of  $x$  and of the parameters  $a$ ,  $b$ , and  $c$  there will certainly be one and but one value of  $y$ . The accuracy of prediction varies with the degree to which it is possible to substitute generalized symbols for myriads of particular cases.

For prediction from plurels consider S. 2, Ch. XII, where the percent of defectives, expected and found, are presented for various combinations of normal and feeble-minded parents and grandparents. A prediction of expected percentages is made on the basis of Mendelian theory and checked against observed percentages to test the fit of this particular application of the theory. Prediction from percentages of persons in plurels is the simplest and most common form of the probability index, which was defined as the ratio of those parties with a characteristic to those with it plus those without it. (Eq. 7, Ch. VI.)

Going on to prediction from the sum of simple probabilities every normal distribution illustrates this. A distribution curve must first be reliably determined. A shoe store, for example, knows the distribution of sizes of feet and orders more medium sizes and fewer extreme sizes in accordance with that curve of past experience. Every stocking-up in any store involves predicting by estimating probabilities either simple, joint, or summed. Since a large number of biological, psychological, and cultural phenomena are normally distributed, the table of the normal probability integral is par excellence a predictive tool.

Every time it is used, as in determining the significance of any index in relation to its standard error of sampling, an implicit prediction is being made, measured in probability terms.

The next three graphs (S. 3, 4, and 5, Ch. XII) illustrate prediction from correlation which is a joint summed probability as explained in Chapter VI. When a specified degree of correlation has been observed between two indices, then, assuming that the same conditions continue, a specified amount of one index predicts specified amounts of the other with specified probability. The detailed mechanisms of such predictions are diagramed with formulae stated in the scattergram S. 1, Ch. VI, and S. 3 and 5, Ch. XII; in the family of ellipses, S. 2, Ch. VI; in the array variances of S. 3, Ch. VI; in the correlation surface of S. 5, Ch. VI; in the overlapping areas of S. 6, Ch. VI; in the percentage of common elements of S. 7, Ch. VI; in the percentage blocks of S. 9, Ch. VI; in the probability whorl of S. 10, Ch. VI and S. 4, Ch. XII; in the cosines and sines of S. 12, Ch. VI; and in the teams of predictors and criteria of S. 22, Ch. VI. Thus in S. 4, Ch. XII the family of probability curves is a new technic, convenient for lay interpretation of prediction probabilities. With these curves reflecting the past experience of that college, the administrator can read off for every Freshman with his known academic record the exact probability of his graduating with any specified rank standing in his class. The probability for an individual is, of course, the same as a percentage for the plurel.<sup>9</sup>

The regression equation and the standard error of estimate of a person's characteristic, estimated from a second known and correlated characteristic, provide another powerful technic for prediction. This technic can be improved by the use of multiple and double multiple correlation (see S. 22, Ch. VI). The index of non-determination (Eq. 14, Ch. XI), as usual, measures the residual unpredictable part of the situation, thus gauging the ignorance of the researcher and challenging to further research.

One powerful predictive application of correlation technics is in correlating indicators of speech with action, indicators of verbally expressed attitudes with behavior. Every one predicts the behavior of other people from their speech, from what they say or write, and how they do it, and most of daily human activity is regulated by such prediction. We plan our appointments and meetings, our transactions, in short our schedule, depending largely on what intentions others have expressed, or we have expressed to them. The enormous mass of such predictions which

come true is forgotten by any one who enlarges on the "unpredictability of human nature." Formal contracts and treaties and published schedules and announcements predict future behavior by speech. If a rigorous count were made of *every* such instance in one's daily life the hypothesis is ventured that well over 90% of speech predictions would come true (within limits of precision, of course, which would have to be specified). The small balance of unfulfilled predictions holds our attention as problems (while habitual routines become unconscious) and creates the popular opinion that prediction of human behavior from speech is unreliable.

The degree of truth in this opinion, however, can be factually measured by correlation technics. By gathering data, *under specified and varied conditions*, the various degrees of correlation of verbal indicators and indicators of other behavior, in short, of correlation between words and deeds, can be scientifically measured, and thereafter prediction of behavior, under similar conditions, can be made from speech with defined degrees of probability and reliability of that probability. (See S. 5, Ch. XII for such validating of a verbal attitude test with marital behavior.)

Any implication in the paragraphs above that S-theory has made possible the predictions, such as reviewed there, is explicitly disclaimed here. Obviously those predictions were made and are being made without any aid from the S-theory. The contribution of S-theory, however, may lie in increasing their (a) precision, and (b) standardization. Predictions are ordinarily made intuitively. Requiring explicit writing of them as in S-theory formulae with observed frequencies of occurrence and non-occurrence, or degrees of occurrence, greatly increases the precision of prediction. This yields a definite index of probability within definite limits of error. Secondly, the S-symbols increase standardization. They provide a single parsimonious set of symbols capable of expressing all statistical processes and formulae, and at the same time expressing many of the conventional and some new sociological concepts. They thus integrate quantitative methodology with the content of Sociology—which have hitherto had little integration. Sociologists have often known little of statistics, and statisticians have often known little of Sociology. But S-symbols provide standard concepts integrating Sociology and Statistics. The

concepts of probability and its compounding are all shown to be combinations of S-concepts of the four sectors, the four scripts, and the simplest operators.

### 3. DYNAMIC PREDICTION

To continue the review of S-theory in connection with the predictive function of science, consider next the examples S. 6, 7, 8, 9, 10, 11, 12, 13, 14, Ch. XII, involving durations, i.e., predicting as dependent on people's ages and the human life cycle. The population pyramid, which S-theory reduces to a formula, predicts longevity and survival (S. 7, Ch. IX), or marital status in the case of S. 6, Ch. XII. Mortality predicted by age and weight in S. 8, fertility predicted by age of parents in S. 9, earnings predicted by age for poor people in S. 7, illustrate the range of S-theory formulae in describing diverse situations involving prediction.

Proceeding with dynamic situations the projection of time series into the future, although attended by the pitfalls of changing conditions, is nevertheless useful within limits. Thus, the velocity trends of institutionalized defectives in S. 18, Ch. XII predicts the need of facilities being developed to care for them in the near future.

The estimates of the U. S. population growth in S. 19, Ch. XII illustrate the use of hypotheses expressed in exact equations, such as the Reed-Pearl logistic curve, as well as forecasts of trends synthesized from alternative conditions (of immigration, birth rates, etc.). The distribution of dynamic phenomena, the amount of an achievement under conditions of specified time limits, illustrates normal probability again in S. 21 and 23, Ch. XII. Forecasting of business trends is verified by comparing a forecast index with the actual later events in S. 17 and 22, Ch. XII. Steady upward trends (Eq. 2, Ch. XI) of social indices in the Soviet Union (S. 23, Ch. XII), regular seasonal cycles (Eq. 3, Ch. XI) of robberies in U. S. cities in S. 25, Ch. XII, and a curved or accelerating trend of income in S. 26, Ch. XII illustrate prediction from trends and cycles. Finally prediction of social control is incipiently illustrated in S. 24, Ch. XII where, by a previously fairly reasonable assumption of static conditions the observed velocity of progress in the interrelations of a maladjusted delin-

quent girl becomes an acceleration from the previous zero velocity. This is a societal force plus interaction of people which defines societal control. The prediction of societal forces in either direction or amount is, of course, difficult proportionately as new conditions may arise constituting additional facilitating, resisting, or redirecting forces for which, until actuarial data can be secured, estimates of probability will be more of an art than a science. This art of estimating from insufficient data is the essence of statesmanship, of "business judgment," of "good management," requiring persons of high intelligence and experience in that field. As data accumulates science replaces art, actuarial calculations replace intuitive judgment, the technician or bureaucrat of inferior ability can satisfactorily and routinely continue the work of a gifted pioneering leader.

### *B. Some Principles for Predicting*

Out of our study of predictive situations some generalizations emerge which are stated below as hypotheses, whose limits and degree of truth, under conditions to be specified, may progressively be determined with further research.

#### 1. DEFINITION OF ACCURACY OF PREDICTION

First let the accuracy of prediction be defined as the degree of agreement between a prediction and its eventuality, between a situation as predicted and that situation as it transpires later on. This degree of agreement may be measured by some form of accuracy index suited to the data, such as a percentage, a goodness of fit probability, a correlation coefficient, etc. Tentatively it looks as if these various forms of probability, from the simple to the complex, which constitute our accuracy index, could be re-expressed as a correlation (including contingency which is correlation of attributes). This suggests a formula (which is here recorded hesitantly because it has been explored very little and not rigorously defined in this volume as yet):

$${}^0: {}^tS \bullet {}^{z:} {}^tS ? = ({}^0: {}^tI) = \text{an index of the accuracy of prediction} \\ \text{(Eq. 2, Ch. XII)}$$

This asserts the amount of correlation between predicted situations (i.e., at dates  ${}^t$  later than the present dates,  ${}^0$ ) and those situations when in the past (i.e., at dates  ${}^t$ , earlier than the

terminal dates, <sup>2</sup>|). Armed with this, or some better index of accuracy of prediction, consider what determines this accuracy.

## 2. CLASSIFICATION

A first principle for increasing the accuracy of predicting is to classify the conditions. The more adequately the relevant conditions have been classified and each class objectively specified, the more accurate the prediction. A corollary is that the more homogeneous the situation predicted, the more accurate the prediction. Homogeneity is achieved by classifying and subclassifying until heterogeneous elements are separated and the final classes are as homogeneous as is possible to achieve, or necessary for a given degree of accuracy.<sup>10</sup> \* Thus, Pemberton found (S. 5, Ch. XI) that the temporal distribution of States adopting compulsory education laws did not fit a normal curve for the United States as a whole, but when the Southern and the Northern States were separately analyzed each was fitted excellently by a normal curve. The negro problem in the South and its absence in the North constituted a heterogeneous variable, and classifying into two classes resulted in each region being homogeneous. If such discoveries are made before the eventuality is completed, i.e., if the good fit of a normal curve had been observed before all States had adopted such laws, prediction of high accuracy could be made as to the annual velocity and final date of States so legislating.

## 3. CALIBRATION

A second principle, or set of principles, for increasing the accuracy of prediction is to calibrate the indices. As elaborated in section IV B of Chapter III, calibration involves determining the limits, the reliability, and the validity. To increase the reliability of the indices, whether they are indices of the situation predicted (i.e., the eventuality) or of the prediction, requires adequate sampling and minimal observer error, observand error, and observee error as measurable by Eqs. 21 and 22, Ch. III. To increase the validity requires higher zero order or multiple correlations between the predictor index and the predicted index, i.e., a higher accuracy of prediction by Eq. 2, Ch. XII. This includes the special case of improving the representativeness of sampling.

\* For Eq. 3, Ch. XII, see notes at end of chapter.

Under this principle of calibration, three corollaries of the sub-principle of multiple correlation should be noted. Multiple correlation between a set of predicting indices and a predicted, or criterion, index will increase according as the predictors are, (a) uncorrelated with each other, (b) highly correlated each with the criterion, and (c) numerous. For, if the predictors are highly intercorrelated, they overlap and each contributes little more than the others towards estimating the criterion. It should be obvious that the optimally weighted sum of a set of indices (which multiple correlation assures) will correlate more highly with the criterion proportionately as each index correlates highly with that criterion. Finally, if the predicting indices are separately valid and are relatively independent of each other, the more of them there are the more completely the criterion will be measured, and hence the more accurately it will be predicted. For examples of these corollaries of the subprinciple of predicting by multiple correlation, study S. 3, Ch. XII, the highly accurate predicting ( $(I, \cdot i) = .98$ ) of the occurrence of notables from environmental characteristics of population density, education, capital, and coolness of climate; or S. 30, Ch. VI, the predicting of academic achievement in college. This last is more fully analyzed in S. 34, Ch. II for another large university, where the relative accuracies of prediction by single predictors and by a multiple set of predictors is graphed as decreasing angles between their vectors. Note that the combination of four college entrance predictors, an English examination, age, rank in secondary school, and College Board examinations has a smaller angle with rank in Freshman year in college, and therefore, predicts it better than any one of the predictors alone. This is mathematically inevitable by the derivation of the formula for multiple correlation. It is evidenced again in that predicting graduation standing from the record of the first three years of college combined is better than from any one-year record alone. (For geometric evidence of these corollaries study S. 22, Ch. VI, the diagram of multiple correlation in  $n$ -space.)

#### 4. STABILITY

A third principle for increasing the accuracy of prediction is to select stable indices, or, alternatively to stabilize the indices.

Stability is measured by the revarying of the index (Eq. 58, Ch. X). Stability is perfect when the sigma measuring revarying is zero and instability increases with that sigma. This is identical with the temporal fluctuations of a static index. (For a discussion of fluctuations of velocities see the section on trends, cycles, and fluctuations in Chapter XI.) Stability is best determined (after removing the trend and cycle), as the residual fluctuations which are individually unpredictable after allowing for the predictable trend and cyclic variation. Obviously the more stable the index to be predicted, the more accurate the prediction is likely to be. If the index is of some phenomena which are subject to human control, it can be stabilized by human purpose and effort. Annual financial expenditures, annual crop and factory productions, and school achievement are a few instances of predictable phenomena where the human intention to reach a certain goal, or not to exceed it, helps to stabilize such indices.

A corollary to this principle of stability is that the accuracy of prediction varies with the immediacy of the eventuality predicted. The more immediate events can be predicted more accurately than the more remote events. This can be proved by the law of joint probability. If the probability of a change in a specified period in a trend, or in any process, or in any static condition, is  $\%I$ , then the probability of the change occurring any time in twice that period is the sum of their separate probabilities, i.e.,  $2(\%I)$ . If  $\%I$ , for example, is 20%, then in five periods the probability of that change is 100%. In general, the probability will cumulate to 100% in  $(100/\%I)$  such periods. Thus, the longer the period, the greater the probability of change in general, and the less accurate the forecast. This is not a rigorous argument and needs qualifying to specified conditions, but it serves to suggest the evidence in favor of this corollary principle that prediction improves with its immediacy.

## 5. INTENSITY OF DESIRE

A fourth generalization, which is here advanced as an hypothesis towards improving the accuracy of participant predictions, deals with the intensity of desire of the participant party for the predicted eventuality (i.e., desideratum). The hypothesis is that

the accuracy of participant prediction correlates positively with the intensity of desire of the participating party.

$$0: \text{I} \bullet D > 0 \quad (\text{Eq. 4, Ch. XII})$$

If the participants care little for the predicted eventuality they will not work for it nor strive to overcome any resisting forces that may arise from sources outside of themselves, i.e., from other parties or from the non-human environment. But if the participants desire the eventuality intensely they will strive mightily for it and have a higher probability of finding or creating some way of overcoming any forces resisting its realization. Since the desires and resulting tensions and efforts of the participants may not be the only forces in the situation, the correlation coefficient expressed in Eq. 4, Ch. XII may be less than unity.

This illustrates again a behavioristic method of dealing with human purposes. Devise calibrated instruments to measure the intensity of a human desire, correlate it with the accuracy of prediction index in a sample of specified situations, and thereafter for comparable situations, the desire is a partial predictor within known limits of probability.

If Sociology is to progress as a science it must improve the accuracy of its predictions (along with its understanding and control of societal phenomena). To improve that accuracy it is necessary to improve the adequate classification, the exact calibration, the stability of the indices recording phenomena, and the intensity of desiring the predicted eventuality. In improving such classification, calibration, stability, and intensity of desire, the symbols of S-theory, with their objective and precise definitions, their parsimony and their systematic comprehensiveness, may prove a useful tool.

#### IV. CONTRIBUTION OF S-THEORY TOWARDS CONTROLLING PHENOMENA

The third and most inclusive aim of science is to increase human control of phenomena. This aim has been little realized by Sociology as yet. There is a vast deal of human control of physical and biological phenomena built up by those sciences, and there is a great deal of human control of societal phenomena, but most of it has grown up in our economic, political, religious,

educational, and other social institutions without the aid of the professional social scientist and his body of systematized knowledge that we call science. This is especially true in the younger sciences. Sociology only attained academic recognition in chairs and courses within the last half century and as a practical profession outside of teaching, as in social work with certificates or Government Civil Service status, etc. within the present century. Most societal control, however, is exercised, as in a Communist State, or a totalitarian Fascist State, or in a democracy at war, by leaders other than professionally trained sociologists.<sup>11</sup>

Even leaving the practice of societal control aside, sociologists have had theories of control in which the central concept lacked a precise objective definition.<sup>12</sup> At this point S-theory may make a contribution with its definition of societal control in the quantic formula,  $T^{-2}IP^2$ , which explicitly denotes an acceleration of a societal change in people, stimulated, or partially caused, by other people. The explicit measuring and recording of the changing characteristic,  $I$ , the time rates,  $T^{-1} : \text{ } T^{-1}$ , and the parties interacting,  ${}^pP_p :: {}^pP_p$  constitute a quantitative definition of societal control. This is the first step in studying control situations and eventually controlling situations. But this rigorous observation has hardly begun among sociologists. In searching all the literature of the collection of 1500 quantitatively recorded situations, although there were many accelerations, and a few forces, and a few interactions, hardly a half dozen situations were found recording their combination in societal control, in even a primitive way.

Thus for instance, S. 22, Ch. XI, a record of the Russian Five-Year Plan estimates and achievements, is an acceleration of production in a plurel and is, therefore, an I-force. But the number of persons in the parties controlling and controlled—the Government, or Communist Party, and the Russian population—are not explicitly recorded, so that the data for calculating indices of societal control, more than a primitive societal force, are not available. The situation implicitly represents societal control but the record falls short of mathematical completeness.

Sociologists speak freely of control and of social planning but very little rigorous data have been gathered recording the parties controlled, the parties controlling, and the nature and amount

and speed of the change constituting the control. Consequently the discussion of this topic is brief here, for although it is the most important topic in all Sociology in the author's opinion, the policy in this volume is to develop theory only as far as facts are available as a basis. Until more quantitative situations involving explicit societal control are recorded in the sociological literature, induction of theory about control will remain meager. Obviously generalization must await the collection of particular instances to generalize. The significance of societal control is the possibility of man's determining his own destiny in the future; its achievement as far as science can contribute depends on the accumulation of more adequate data, so that from particular truths in this field more general truth can be induced. This is the central problem of Sociology.

#### V. *S-SITUATIONS*

The new formulae dealing with prediction in this Chapter were induced from S-situations such as the following. This sample has been arranged according to their quantic numbers to follow the sequence of Chapters in this volume. While they are only a suggestion of the extent to which predicting is going on at present, they illustrate possibilities (for further refinement) of the social scientist's ability to predict.

## S. 1

## LEVELS IN THE PROCESS OF SYMBOLIC SUBSTITUTION

(Read up)

<i>Level of Experience</i>	<i>Description of Experience</i>	<i>Form of Symbolic Substitution</i>	<i>Type of Educational Transmission</i>	<i>Level of Trial and Error Process</i>
6th level	$y = a + bx + cx^2$	Abstract mathematical formula or family of curve	Highly abstract symbolic substitution	
5th level	$y = 100 - 3.5t + 0.264t^2$ (equation fitted to the trend of the tabulated figures)	Specific arithmetical and algebraic symbols in an equation that summarizes a trend of the data	Specific numerical substitutes	Symbols substituted in the process of mental manipulation
4th level	Graphs based upon tabulated numbers of cases per month	Generalized description in graphs and tabulations		
3rd level	Monthly statistical reports of cases of a poor relief department	Specific numerical data	Summarized description in number symbols	Numerical substitute stimuli and responses
2nd level	Case records of names, addresses, notes on needs, resources, etc., of relief clients	Specific records of human experience stored in verbal symbols	Qualitative word-symbol substitutes for non-present persons	Qualitative word-symbol substitutes for direct experience
1st level	Memories of sense impressions (after office hours) in the minds of social workers who interviewed clients	Visual and auditory images as substitutes for experience	Conversation and discussion	
0 level	Unemployed persons in need of relief	Sense perceptions	Personal face-to-face contacts	Overt level of wasteful manual manipulation

Ref.: Chapin, F. Stuart, *Contemporary American Institutions*, Harper and Brothers, 1935, p. 140.

Descriptive formula:  $S_1 = {}^1I^{+1} :: I_1^0$

Legend:

$S_1$  = The situation

records

${}^1I^{+1}$  = 6 levels of experience

:: = cross-classified

Quantic number = 0;1;0;0

with

$I_1^0$  = 4 categories of symbolic substitution

## S. 2

No.	Matings	Percent of Expected	Defectives Found
1	NN × NN	0	0
2	NN × Nn	0	0
3	NN × nn	0	37.5
4	Nn × Nn	25	33.2
5	Nn × nn	50	53.6
6	nn × nn	100	77.3
N = Normal and n = Feeble-minded.			

Ref.: Bushee, Frederick A., *Principles of Sociology*, Henry Holt and Co., 1923, p. 347.

*Descriptive formula:*  $S_2 = \%P_p : q$

*Quantic number* = 0;0;0;1

*Legend:*

$S_2$  = The situation

: = subclassified into

records

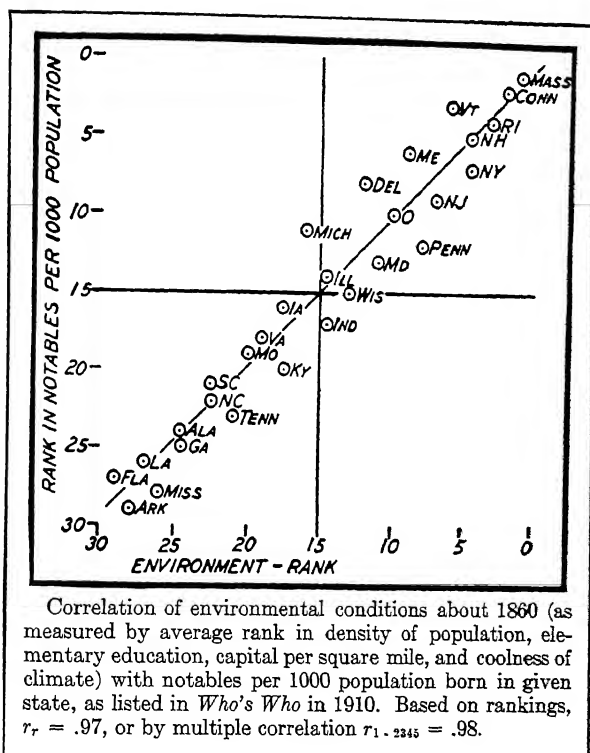
$q = 2$  probability  
plurels  $\left\{ \begin{array}{l} \text{"expected"} \\ \text{and} \\ \text{"found"} \end{array} \right.$

$\%P$  = the percent of a population  
studied

in each of

$|_p = 6$  Mendelian plurels

## S. 3



Ref.: Reinhardt, J., and Davis, G., *Principles and Methods of Sociology*, Prentice-Hall, 1932, p. 262.

Descriptive formula:  $S_3 = {}^tT^0 : {}^i(I) :: {}^M(I_i) : \underline{P}_p$

$$r_{IJ} = (\sigma I_{I,J}) = .98$$

Quantic number = 0,2;0;1

Legend:

$S_3$  = The situation

records for

${}^tT^0$  = 2 dates

$::$  = a cross-classification of

(I) =  $\%P$  = an index of notables per 1000 population

${}^i|$  = in rank units

and

${}^M(I)$  = a mean index of

$|_i$  = 4 environmental indices

with corresponding

$\underline{P}_p$  = frequency of identified States.

$r_{IJ}$  = The correlation coefficient which is also

$(\sigma I_{I,J})$  = a scalar product in sigma units

= .98

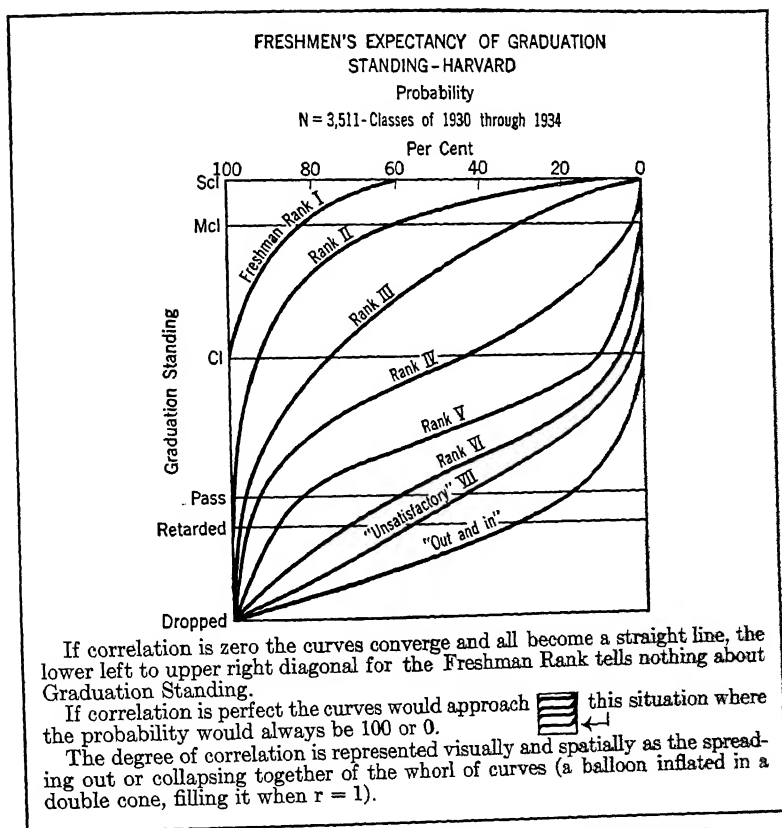
S. 3 (*Continued*)*Comment:*

Even with the unusually high correlation of .98 the percentage of undetermined elements ( $100k^2$ ) is 4%. Prediction of notables emerging may be made from a knowledge of these environmental conditions with 96% of determinateness ( $100r^2$ ). Since the correlation and antecedence criteria of causation are here fulfilled, there is evidence that under the conditions of these data, whatever these 4 environmental indices measure are 96% of the causation of notables. The deeper question is to what extent this percent of causation would vary as the conditions might vary. To answer this requires similar data gathered for other conditions (of cultures, periods, etc.).

Thus, are these environmental conditions directly causative or partly selective? That is, if such an environment were created for a given population, would more notables result? Or conversely—to what extent does superior heredity result in creating a superior environment, and the combination then show more notables emerging?

Thus prediction is possible from known correlation; but without fuller knowledge of causation, duplicating the later index by creating the antecedent index may or may not be possible.

## S. 4



Ref.: Dodd, S. C., Report to the Dean of Harvard College, 1935 (unpublished).

Descriptive formula:  $S_4 = {}^iI :: {}^iI : \%P$

Legend:

$S_4$  = The situation  
records for every value of  
 ${}^iI$  = graduation standing (in ordinal  
units)  
 $::$  = cross-classified with

Quantic number = 0.2;0;1

${}^iI$  = Freshman standing (in ordinal  
units)  
: = a corresponding  
 $\%P$  = population percentage or prob-  
ability of attainment

Comment:

This graph illustrates a new method of portraying correlation in terms of a family of probability curves which can be readily put to use by laymen for the prediction, based on previous experience, of individual cases with known probability.

## S. 5

PERCENTAGE DISTRIBUTION OF PREDICTION SCORES FOR THOSE WHO ARE DIVORCED, SEPARATED, HAVE CONTEMPLATED DIVORCE OR SEPARATION AND HAVE NOT CONTEMPLATED DIVORCE OR SEPARATION

Prediction Score	Marital Status				Number of Cases
	Divorced	Separated	Have Contemplated Divorce or Separation	Have Not Contemplated Divorce or Separation	
700-789	0.0	0.0	9.1	90.9	11
620-699	2.9	0.0	5.9	91.2	68
540-619	2.9	4.3	6.5	86.3	139
460-539	13.9	15.0	13.9	57.2	173
380-459	25.0	17.0	16.0	42.0	100
300-379	34.2	21.9	21.9	21.9	41
220-299	50.0	37.5	12.5	0.0	8
Number cases	73	61	64	342	540 *

\* Fourteen cases were added to the original sample of 526.

Ref.: Burgess, Ernest W., and Cottrell, Leonard S., "The Prediction of Adjustment in Marriage," *American Sociological Review*, Vol. I, No. 5, Oct., 1936, p. 751.

Descriptive formula:  $S_5 = {}_1I :: {}^1I : \%P$

Quantic number = 0;2;0;1

Legend:

$S_5$  = The situation

${}^1I$  = in 4 ordinal degrees

records

$::$  = with their corresponding

$I$  = a prediction indicant

$\%P$  = percentage frequencies of persons

${}_1I$  = in 7 class-intervals

$::$  = cross-classified with

$I$  = a marital status indicant

Comment:

This correlation scattergram is typical of the process whereby prediction of human phenomena is improved. A rough but socially accepted scale of ordinal categories is taken as the criterion—in this case, the criterion of marital adjustment. Then an attitude, or other type of behavioristic scale, is constructed with cardinal units which discriminate degrees of marital adjustment more finely. This attitude has the two additional merits that it can be determined in a few minutes instead of requiring the period of years which the criterion takes to be determined, and that it can be determined in advance and hence in time to modify or prevent any undesired predicted outcome. In proportion to the percentage of determination ( $100r^2$ ) of the predictor scale and the criterion, the predictor can be used to determine in advance, within known limits of error, the socially significant criterial situation.

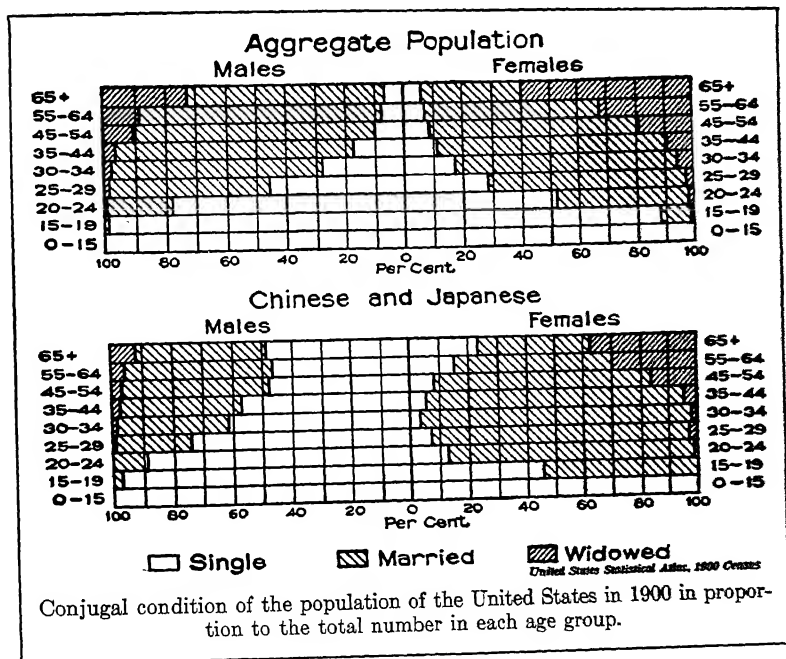
In this situation the criterion, marital adjustment, was reduced to an indicant and correlated with the "prediction score." The correlation was .51 on one sample ( $P = 526$ ) and .48 on another checking sample ( $P = 155$ ). This means

## S. 5 (Continued)

a percentage of determination of about 25%. This may seem low to some students, but it must be remembered that:

- it is probably better than previously existing technics (most of which have never been quantitatively calibrated), and
- it can be further improved, as the authors say, by refinements in the instruments.

## S. 6



Ref.: Brinton, Willard C., *Graphic Methods for Presenting Facts*, Engineering Magazine Co., 1923, p. 168.

S. 6 (*Continued*)

*Descriptive formula:*  $S_6 = {}^tT^0 : ({}_tT^{+1} : \%P_p : q)_r$

*Quantic number* = 1;0;0;1

*Legend:*

$S_6$ = The situation	$ _p$ = 2 sex plurels
records	each subdivided into
${}^tT^0$ = in 1900	$ _q$ = 3 conjugal plurels
for each of	and all repeated for
${}_tT^{+1}$ = 9 unequal age groups	each of
${}_t^*$ = with terminal ages stated	$ _r$ = 2 racial plurels
$\%P$ = a population percentage	
for each of	

*Comment:*

A simple form of actuarial prediction is illustrated here. Based on the experience of the population represented in this situation, the probability of a person's being in a given "conjugal condition" at any age may be determined and insurance, housing plans, and other societal components may be more intelligently planned as far as affected by the conjugal status of the population.

Thus, for example, the probability of a Chinese or Japanese woman being a widow after the age of 65 is:

$${}_{t>65} : \%P_{p''q''r''} = 38\%$$

or of any man aged 30-34 being single is:

$$({}_{t'=30-34}) \%P_{p'q'r'} = 28\%$$

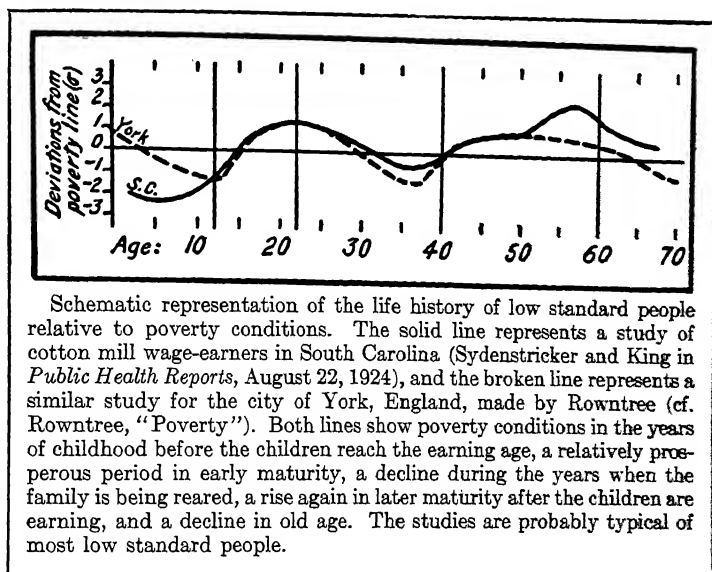
Comparisons of probabilities can be expressed in equations and compounded in various ways. Additional data can result in joint probabilities or subclasses with more exact prediction. Or aggregates of classes may be compared as in the matrix equation:

$$({}_tT^{+1} : \%P_{p''q''r''>r'})$$

which states that at all ages the corresponding percentages of widowed females in the general population exceed those percentages among the Chinese and Japanese.

Trends can be summarized as in stating in the notation above the obvious curvilinear correlation between age and conjugal condition, either by the equation of some fitted curve, or as a summarizing correlation ratio.

## S. 7



Schematic representation of the life history of low standard people relative to poverty conditions. The solid line represents a study of cotton mill wage-earners in South Carolina (Sydenstricker and King in *Public Health Reports*, August 22, 1924), and the broken line represents a similar study for the city of York, England, made by Rowntree (cf. Rowntree, "Poverty"). Both lines show poverty conditions in the years of childhood before the children reach the earning age, a relatively prosperous period in early maturity, a decline during the years when the family is being reared, a rise again in later maturity after the children are earning, and a decline in old age. The studies are probably typical of most low standard people.

Ref.: Reinhardt J., and Davis, G., *Principles and Methods of Sociology*, Prentice-Hall, 1932, p. 586.

Descriptive formula:  $S_7 = {}_tT^{+1} : \underline{P}_p : {}_\sigma I$

Quantic number = 1;1;0;1

Legend:

$S_7$  = The situation

records for each of

${}_tT$  = the 70 years of life

for each of

$\underline{P}_p$  = 2 plurels  $\left\{ \begin{array}{l} \text{in Yorkshire and} \\ \text{in South Carolina} \end{array} \right.$

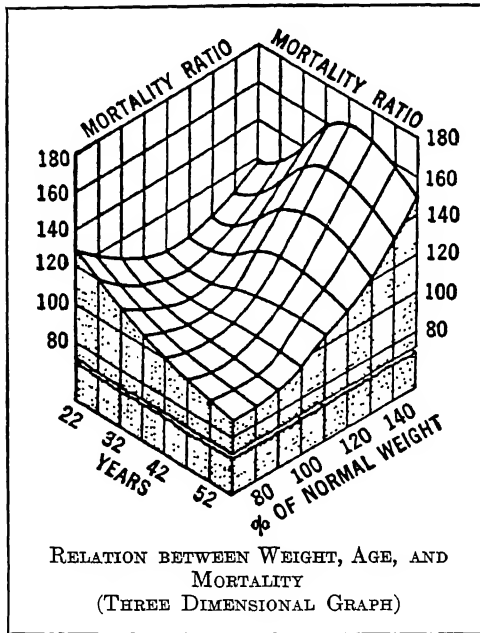
$I$  = an indicant of poverty

$\sigma$  = in standard deviation units

Comment:

This situation illustrates prediction from a type. The typical life-cycle of low income earners is graphed as a type, or average, around which other persons may be expected to deviate. The extent of deviations might be shown by parallel curves in sigma or percentile units, thus enabling more precise prediction of an expected income at each age and a specified probability of any specified deviation from the mean.

## S. 8



*Ref.: Court, A. T., "Measuring Joint Causation,"  
Journal of the American Statistical Association, Sept.,  
1930.*

*Descriptive formula:*  $S_8 = {}^tT^{+1} :: \%I : \%P$   
*Legend:*

*Quantic number* = 1;1;0;1

$S_8$  = The situation  
records for each of

with

${}^tT^{+1}$  = 35 ages

$\%I$  = the % of normal weight

'| = beginning at 22

: = a corresponding

:: = cross-classified

$\%P$  = percentage of deaths

*Comment:*

Prediction from joint probability of two simultaneous predictors (age and weight) is illustrated.

The possibility of predicting even when cause-effect relations are unknown is evidenced here. The age and weight indices are antecedent to and correlated with mortality and are to that extent causes of mortality. The conditions under which this is true and which, if varied, might invalidate the probabilities are not explicitly recorded in the graph.

## S. 9

PREDICTED FERTILITY OF MARRIAGE FOR SELECTED AGES OF  
HUSBAND AND WIFE

Age of Husband	17	20	23	26	Age of Wife	29	32	35	38	41
17	9.58	8.24	6.85	5.41	3.92	2.38	0.79			
20	9.31	8.09	6.82	5.49	4.12	2.69	1.20			
23	8.99	7.88	6.72	5.51	4.25	2.93	1.57	0.15		
26	8.60	7.61	6.56	5.46	4.31	3.11	1.86	0.56		
29	8.15	7.27	6.34	5.35	4.32	3.23	2.10	0.91		
32	7.63	6.87	6.05	5.18	4.26	3.29	2.27	1.19	0.07	
35	7.06	6.40	5.70	4.95	4.14	3.28	2.37	1.41	0.40	
38	6.42	5.88	5.29	4.65	3.96	3.21	2.42	1.57	0.67	
41	5.71	5.29	4.81	4.29	3.71	3.08	2.39	1.67	0.88	
44	4.95	4.64	4.27	3.86	3.40	2.88	2.32	1.70	1.03	
47	4.12	3.92	3.67	3.38	3.03	2.63	2.17	1.67	1.11	
50	3.23	3.14	3.01	2.83	2.59	2.30	1.97	1.58	1.14	

Ref.: Bushee, Frederick A., *Principles of Sociology*, Henry Holt and Co., 1923, p. 396.

Descriptive formula:  $S_9 = {}_tT^{+1} :: {}_uT^{+1} : P$

Quantic number = 2;0;0;1

Legend:

$S_9$  = The situation

$::$  = cross-classified with

records in terms of

${}_u$  = 9 such age classes of wives

$T$  = 3-year age classes

with corresponding

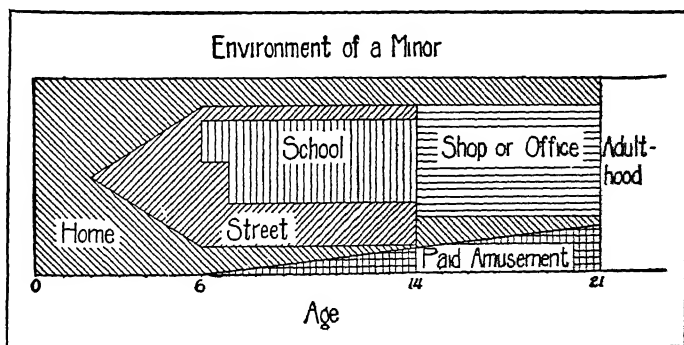
${}_t$  = 12 such age classes of husbands

$P$  = predicted fertility

Comment:

Obviously  $|{}_uT \cdot P| > |{}_tT \cdot P| \neq 0$ , i.e., age and fertility are negatively correlated for each sex. The multiple correlation of the ages of both parents with fertility is even higher, permitting more accurate prediction of offspring than is possible from knowing one parent's age alone.

## S. 10



Ref.: Kelly, Truman L., *Statistical Method*, Macmillan, 1923, p. 38.

S. 10 (*Continued*)

*Descriptive formula:*  $S_{10} = {}^4P : {}_tT^{+1} : \underline{T}_t^{+1}$

*Quantic number* = 2;0;0;1

*Legend:*

$S_{10}$  = The situation

${}_tT^{+1}$  = 21 ages

records for

$\underline{T}_t^{+1}$  = 5 kinds of environmental time expenditure

${}^4P$  = a typical minor

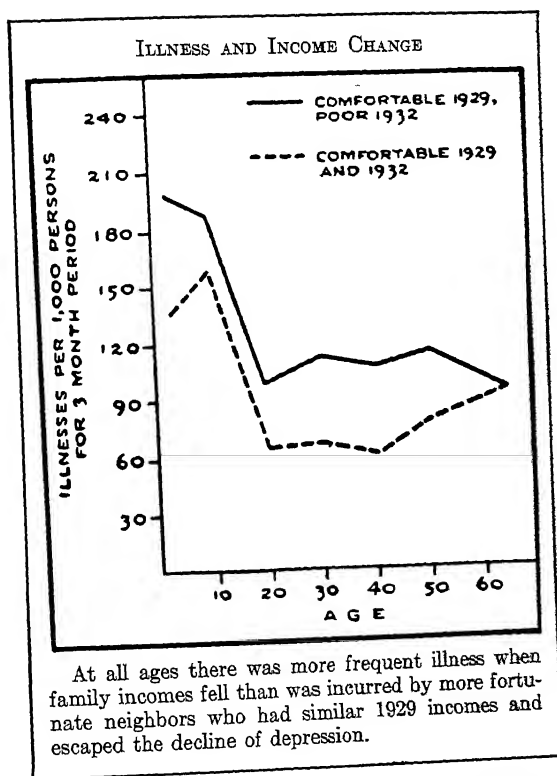
for each of

*Comment:*

Prediction here is from a typical temporal pattern, a daily cycle at each age of a minor. It is an aggregation of implicit means of time expenditure which are norms, or the "most probable" patterns. Note that the formula for a cycle (Eq. 3, Ch. XI) holds, in that at every age the correlation between the hour of the day and the type of environment would be measured by a high coefficient of contingency.

The hours spent in the five environments are an implicit attribute-time product which, if written explicitly (in order to calculate a coefficient of contingency), would read:  $T_t^{+1} = I_1^0 : T_0^{+1}$ .

## S. 11



Ref.: Sydenstricker, Edgar T., "Sickness and the New Poor,"  
*Survey Graphic*, Vol. XXIII, No. 4, April, 1934, p. 162.

Descriptive formula:  $S_{11} = {}^tT^{+1} : ({}_pPT^{-1})_p$

Quantic number = 19;0;0;1

Legend:

$S_{11}$  = The situation

records for each of

$t|$  = 7

$T^{+1}$  = 10-year age classes

$'|$  = beginning at 2 years  
 for each of

$|_p$  = 2 wealth pleurels

$()$  = an index

of

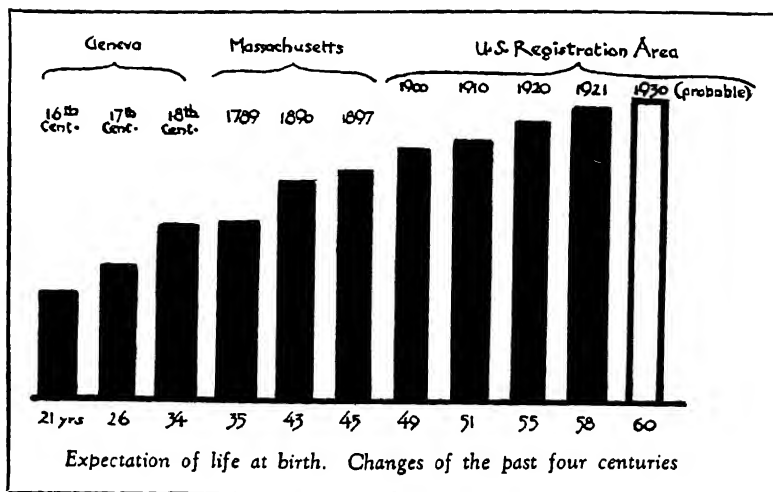
$\%P$  = persons sick per 1000 popula-  
 tion

$T^{-1}$  = per 3-month period

Comment:

The prediction here is by classes, showing differential morbidity as probable for rich and poor at every age up to 60.

## S. 12



Ref.: Emerson, Haven, "Are We Fostering the Unfit," *Survey*, Vol. LXVI, No. 1., April, 1931, p. 51.

Descriptive formula:  $S_{12} = \underline{P}_D : {}^tT^{-1} : T^{+1}$

Legend:

$S_{12}$  = The situation  
records for each of  
 $\underline{P}_D$  = 3 geographic purels  
for each of

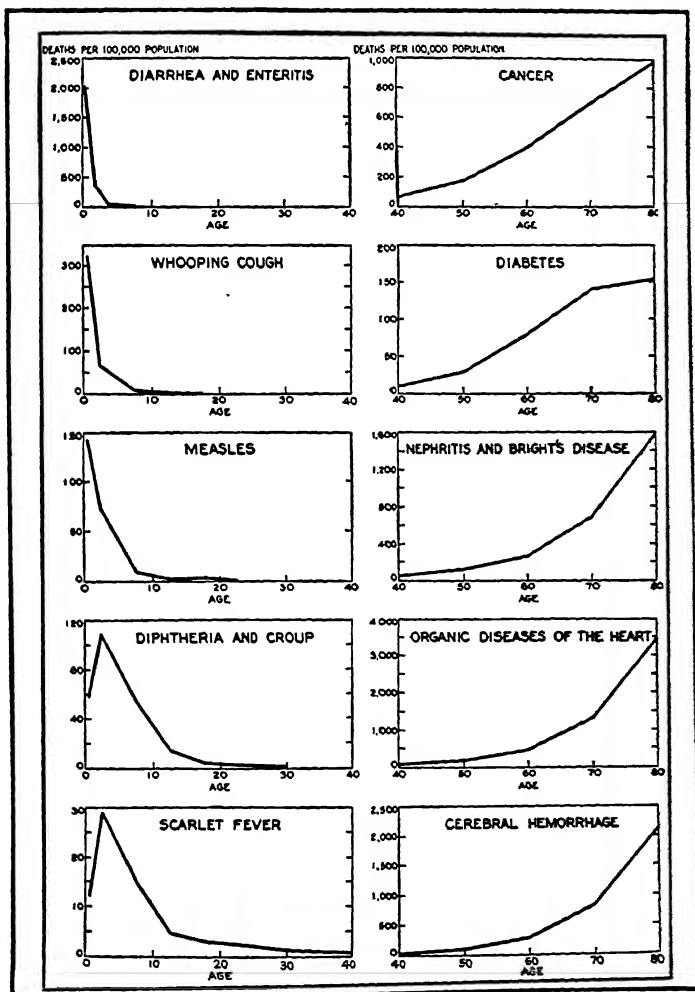
Quantic number = 91;0;0;1

${}^tT^{-1}$  = a series of 11 dates  
the corresponding  
 $T^{+1}$  = longevity

Comment:

The average velocity for the last four centuries of this during process has been approximately .1 year per year, i.e., expectation of life has been increased one year in each decade. Proportionally as similar causes may continue to operate, the trend may be projected into the near future to predict further increase of average longevity at this velocity.

## S. 13



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, Vol. I, 1933, p. 614.

Descriptive formula:  $S_{13} = ({}_tT^{+1} : \%P, T^{-1})_p$

Legend:

$S_{13}$  = The situation  
records for each of  
 $|_p$  = 10 disease plurels  
for each of

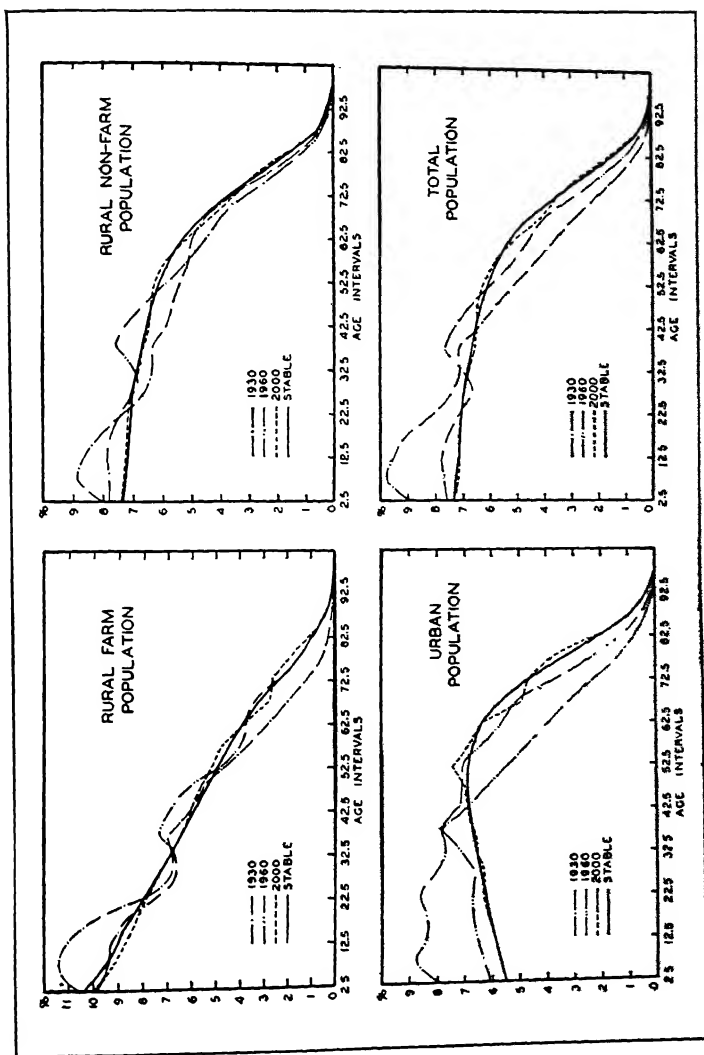
Quantic number = 19;0;0;1

${}_tT^{+1}$  = 8 10-year age classes  
 $T^{-1}$  = an annual rate in 1920  
of  
 $\%P$  = persons dying per 100,000  
population

S. 13 (*Continued*)*Comment:*

Differential probabilities for different diseases and for every age are here given separately. These constitute a partial analysis of the causes of total mortality. Prediction of progress in a summary index, such as decreasing mortality, can be made more accurate by the knowledge of the trend of the component indices of specific mortalities. (See discussion of Eq. 11, Ch. X.)

S. 14



Ref.: Karpinos, Bernard D., "Length of Time Required for the Stabilization of a Population," *Amer. Journ. Soc.*, Vol. XLII, No. 4, Jan., 1931, p. 510.

S. 14 (*Continued*)

*Descriptive formula:*  $S_{14} = {}^tT^{-1} : ({}_tT^{+1} : \%P)_{p, \Sigma p}$       *Quantic number* = 91;0;0;1

*Legend:*

$S_{14}$  = The situation

for each of

records for each of

${}_tT^{+1}$  = ages 1-100

$|_{p, \Sigma p}$  = 3 density plurels and their  
total

the corresponding

for each of

$\%P$  = percentage frequencies of  
people

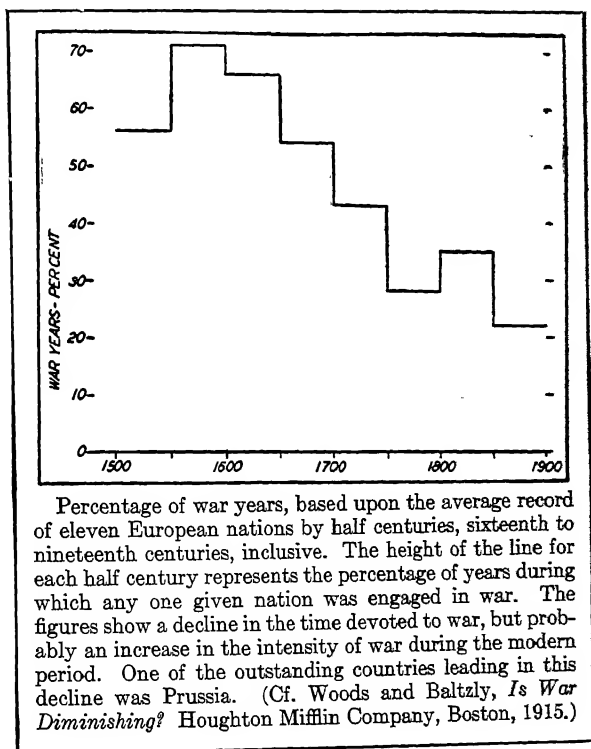
${}^tT^{-1}$  = 4 dates

*Comment:*

Note the waves in the curve (like ripples from a stone thrown into a pool). The crest of the wave starts in children (age  $10 \pm$ ) in 1930, reaches adulthood (age  $40 \pm$ ) by 1960, is waning away around senility (age  $75 \pm$ ) by 2000 A.D. and disappears in the stable curve thereafter. This "pure" wave is seen most clearly in the rural farm curve. It is complicated as if from cross ripples (births and immigration?) in the urban population.

Here are definite predictions of the future, founded on explicit hypotheses, awaiting crucial testing by the censuses of the decades ahead.

## S. 15



Ref.: Reinhardt, J., and Davis, G., *Principles and Methods of Sociology*, Prentice-Hall, 1932, p. 246.

Descriptive formula:  $S_{15} = \underline{P}_{\Sigma p} : {}_tT^{-1} : \%T^{+1}$

Quantic number = 91;0;0;1

Legend:

$S_{15}$  = The situation  
records for

$\underline{P}_{\Sigma p}$  = a plural of 11 nations  
for each of

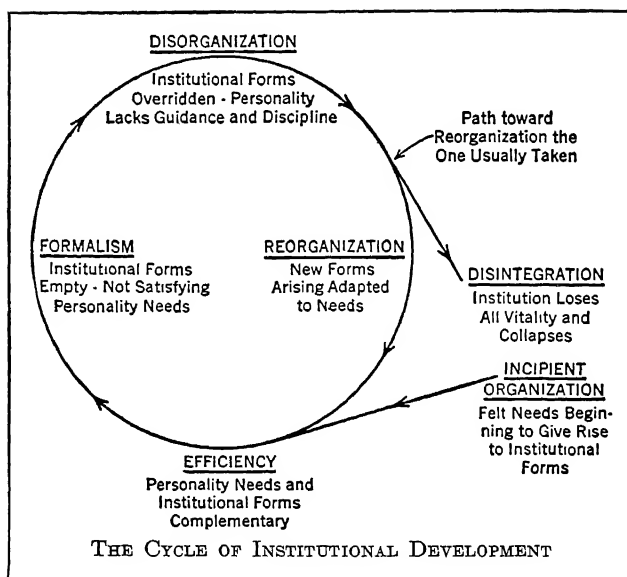
${}_tT^{-1}$  = 8 half centuries  
'| = beginning in 1500 A.D.  
the

$\%T^{+1}$  = % of years of war

Comment:

Prediction by projecting a past trend into the future assumes continuance of similar conditions. Appraisal of such similarity is the subjective element in the prediction. Changed conditions may at any date alter the trend. Analysis of component trends suggests that wars of prolonged durations become less probable in proportion to their increasing severity with modern organization and technical knowledge.

## S. 16



Ref.: Cooley, Angell and Carr, *Introductory Sociology*, Scribners, 1933, p. 407.

Descriptive formula:  $S_{16} = {}_tT^{-1} : I^0$

Legend:

$S_{16}$  = The situation  
records for each of  
 $t$  = 6 periods

Quantic number = 9;0;0;0

of  
 $T^{-1}$  = unstated length  
 $I^0$  = an attribute of institutional development

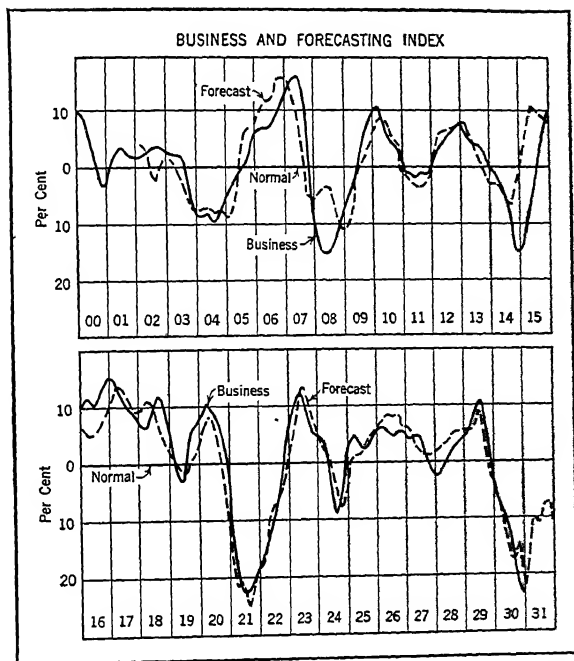
Comment:

This diagram constitutes a hypothetical cycle of institutional development. It is a suggestive hypothesis and can be more exactly tested in proportion to the improvement of indicators of institutional forms and personality needs. Such indicators should identify the stage in the cycle at which any given institution is at the moment, and thus facilitate prediction of the next stages.

The six institutional stages are perfectly correlated with the successive periods by definition, since the ordinal stages define the periods of time. This illustrates our formula for a cycle (Eq. 3, Ch. XI) in the case of a velocity instead of a celeration:

$${}_t(T^{+1} \cdot {}_tI^{+1} = 1.0)$$

## S. 17



Ref.: Smith, Bradford B., "A Forecasting Index for Business," *Journal of The American Statistical Association*, Vol. XXVI, No. 174, June, 1931, p. 126.

Descriptive formula:  $S_{17} = {}^1T^{-1} : I_i$

Legend:

$S_{17}$  = The situation  
records for each of

${}^1T^{-1}$  = 31 years

Quantic number = 9;1;0;0

${}^1$  = beginning in 1900

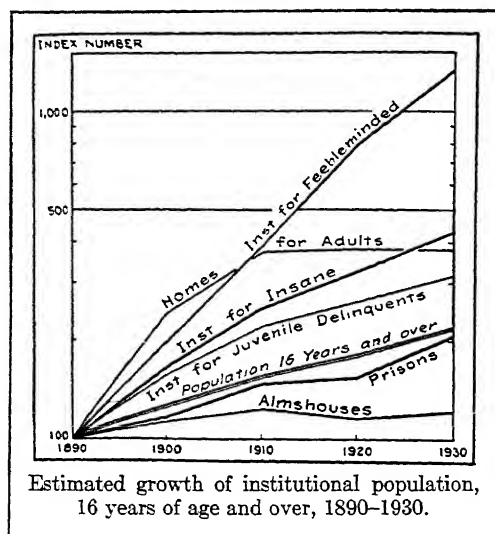
$I_i$  = 2 indicants  
of business

{ "forecast"  
and  
"normal"

Comment:

Here a technic of prediction, a forecasting index, is tested by comparison with actual events. If the correlation is unity ( $I_i \cdot I_{ii} = 1.0$ ), the forecast is perfect.

## S. 18



Ref.: President's Research Committee, *Recent Social Trends*, McGraw-Hill, Vol. I, 1933, p. 306.

Descriptive formula:  $S_{18} = {}^tT^{-1} : \%P_p$

Quantic number = 9;0;0;1

Legend:

$S_{18}$  = The situation

$\%P$  = a population index

records for each of

for each of

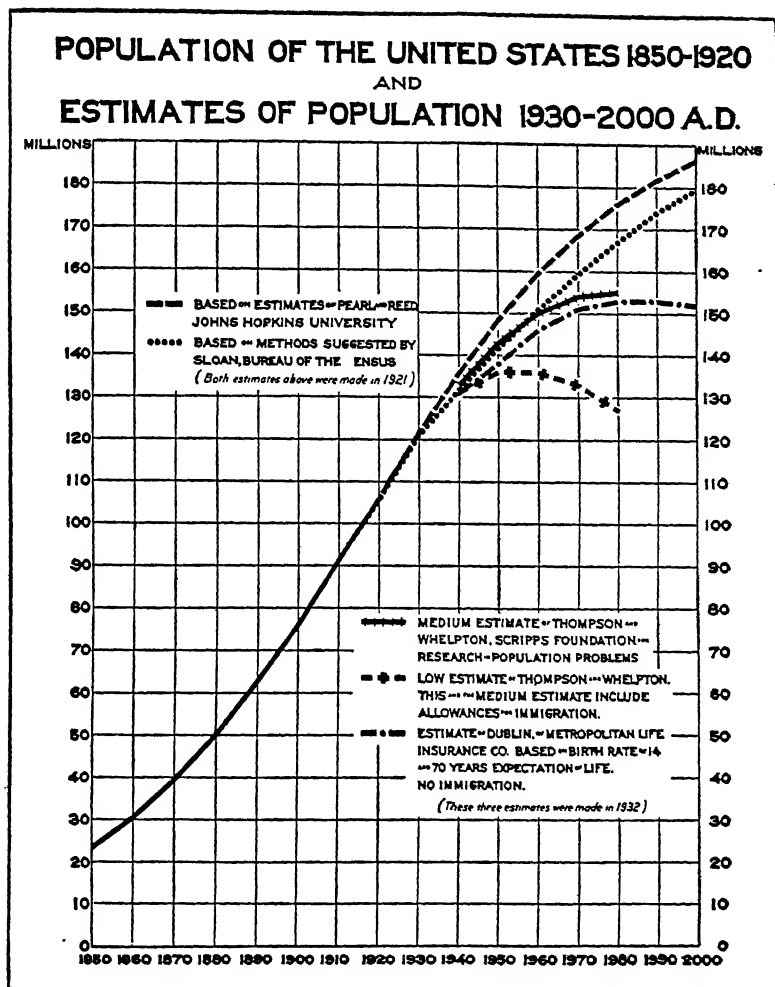
${}^tT^{-1}$  = 5 decennial censuses 1890-1930

$|_p$  = 1 general and 6 institutional plurels

Comment:

Prediction from trends of comparable index numbers here has a criterion for comparison in the trend of the adult population. Prisons parallel the population growth, almshouses lag, while the other four institutions lead. A straight line trend in this graph on a semi-logarithmic scale means a constant *percentage* of annual growth, not a constant absolute increment. The growth of institutionalized population does not accurately measure the growth of this social problem plurel, since the boundary defining the minimal who are sent to an institution may be rising. Borderline persons, formerly left at large, are increasingly sent to an institution in recent years.

## S. 19



POPULATION OF THE UNITED STATES, 1850-1930, AND ESTIMATES FOR 1930-2000

The earlier estimates by Pearl and Reed and by Sloane represent projections into the future of curves fitted to gross population growth before 1920. Later estimates indicate an earlier cessation of growth with a lower maximum population. The estimate by Dublin, 1932, agrees closely with the "medium" estimate by Thompson and Whelpton. The latter authors now regard this "medium" estimate as too high.

S. 19 (*Continued*)*Descriptive formula:*  $S_{19} = {}^tT^{-1} : P_p$ *Quantic number* = 9;0;0;1*Legend:* $S_{19}$  = The situation

P = the United States population

records for each of

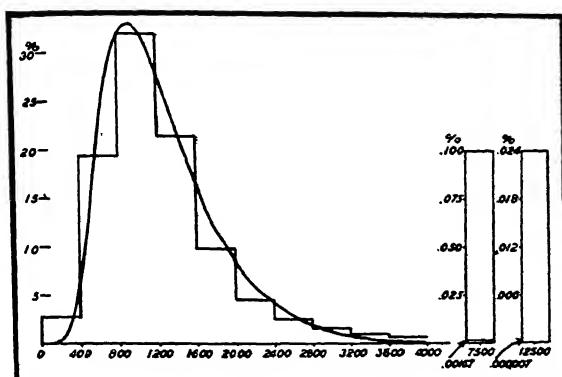
in each of

 ${}^tT^{-1}$  = 16 decennial censuses $|_p$  = 5 classes, or bases of estimate*Comment:*

Two principles of prediction are illustrated in these forecasts of population growth. The first is that the accuracy of prediction varies inversely with the immediacy of the events predicted. The more remote the event the less accurate the prediction. The earlier estimates of 1921 may be expected to appear by the test of the 1940 census less accurate than the estimates of 1932.

The second principle is that prediction varies with its components—here these are birth rate, longevity, and net immigration. Accordingly as these components are differently estimated, their resultants will vary.

## S. 20



Distribution of income in the United States in 1918, according to the National Bureau of Economic Research, and logarithmic normal curve of distribution fitted to the first and third quartiles of the data. The distribution differs from the usual norm in that the data above approximately thirty-five hundred dollars run materially above the computed normal as indicated by the magnified cross sections of the curve at \$7500 and \$12,500, respectively. In the former case the actual distribution is .100%, while the normal is only .00167%. At the latter point the actual distribution is .024%, while the normal is practically 0. The normal curve, as a logarithmic point binomial, would end at a point below \$25,000. For data see Table 33.

Income Class	Mid-point <i>m</i>	Income Recipients ( <i>f</i> %)	Normal Curve (Ht. at <i>m</i> )
0- \$ 400	\$ 200	2.84	.18
400- 800	600	19.51	22.06
800- 1200	1000	32.16	32.46
1200- 1600	1400	21.53	22.07
1600- 2000	1800	9.88	11.87
2000- 2400	2200	4.64	5.90
2400- 2800	2600	2.63	2.87
2800- 3200	3000	1.61	1.40
3200- 3600	3400	1.07	.69
3600- 4000	3800	.74	.35
.....	.....	.....	.....
7300- 7700	7500	.10	.00167
.....	.....	.....	.....
12300-12700	12500	.024	.000007
.....	.....	.....	.....
17300-17700	17500	.01	.....

Percentage distribution of income in the United States, 1918 (National Bureau of Economic Research), and normal curve fitted to first and third quartiles (\$833 and \$1574). Normal curve taken from a study by Peterson and Westmeyer.

S. 20 (*Continued*)

*Descriptive formula:*  $S_{20} = ({}_1(IT^{-1}) : \%P)_p$

*Quantic number* = 9 ; 1 ; 0 ; 1

*Legend:*

$S_{20}$  = The situation

the corresponding

records for each of

$\%P$  = % of recipients

${}_1|$  = 13 class-intervals

in each of

of

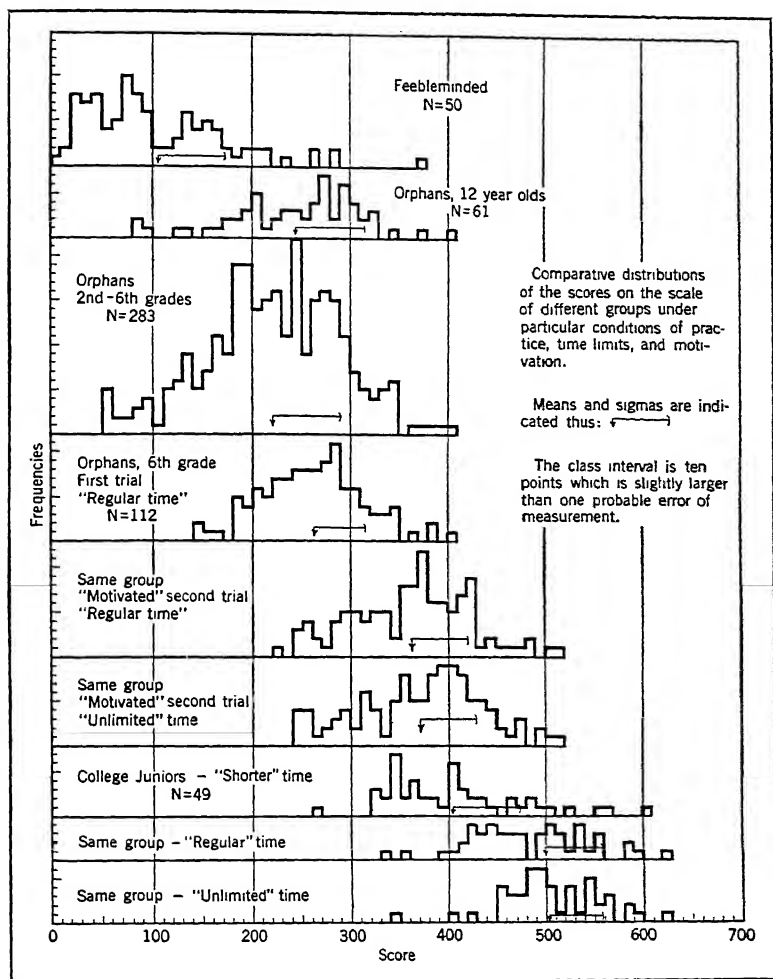
$|_p$  = 2 plurels  $\left\{ \begin{array}{l} \text{an actual and a} \\ \text{normally dis-} \\ \text{tributed} \end{array} \right.$

$IT^{-1}$  = annual income

*Comment:*

The usual positive skewness of an income curve may be noted in this graph in connection with the trends of capitalism and communism, as defined by the statistical moments in Chapter V.

## S. 21



S. 21 (*Continued*)

*Descriptive formula:*  $S_{21} = {}^tT^{-1} : {}_p(M, \sigma, I : P)$

*Quantic number* = 9;1;0;1

*Legend:*

$S_{21}$  = The situation  
records for each of

${}_iI$  = an intelligence score in 65 PE  
units

${}^tT^{-1}$  = several periods and occasions  
of testing

with corresponding

$P$  = frequencies of persons

for

and with

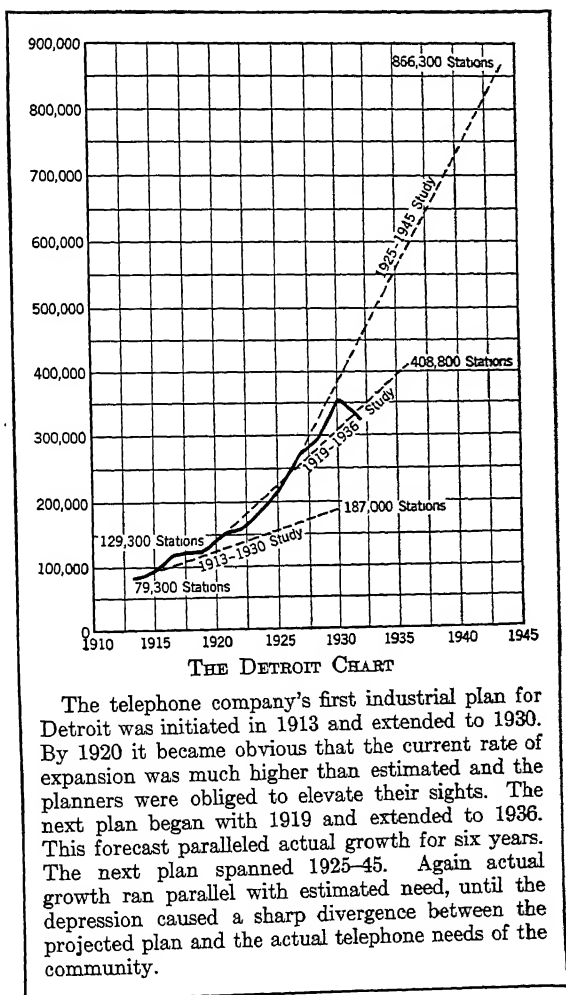
${}_p|$  = 5 plurels

${}^{M, \sigma, |}$  = means and sigmas also stated

*Comment:*

Distribution curves form a major technic of prediction in the social sciences. The most probable value for any future case is near the mean: the probability of any specified deviation from the mean is given by the corresponding area that is cut off under the curve.

## S. 22



Ref.: Lindeman, E. C., "Telephones—Forecasting in Public Service," *Survey*, Vol. LXVII, No. 11, March 1, 1932, p. 598.

S. 22 (*Continued*)

*Descriptive formula:*  $S_{22} = \underline{P}_i : t|T^{-1} : I_i$

*Quantic number* = 9;1;0;1

*Legend:*

$S_{22}$  = The situation  
records for

$\underline{P}_i$  = Detroit  
for each of

$t|$  = 13 periods  
each of

$T^{-1}$  = 25 years

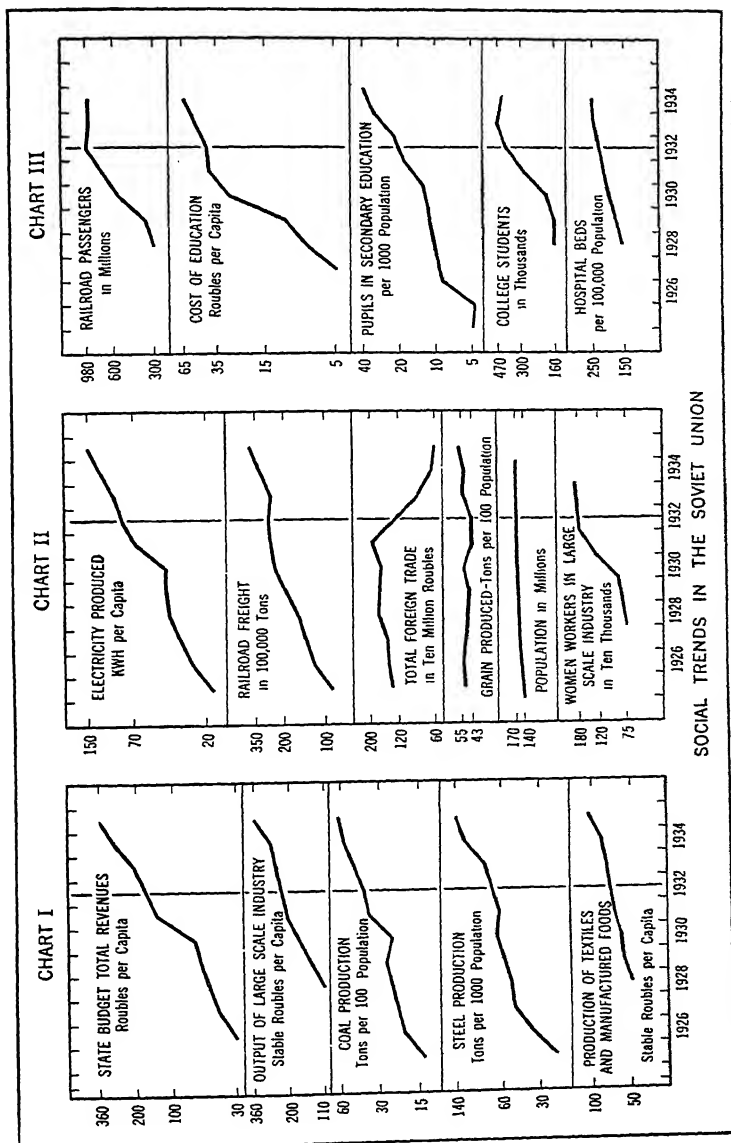
$'|$  = beginning in 1913,

$I$  = the number of telephones  
determined by

$|_i$  = 4 indicants (namely 3 forecasts and the actual)

*Comment:*

Prediction here is a commercial necessity. Periodic revision increases its accuracy. Conservatism leans towards projecting the trend of velocity rather than the trend of acceleration in making these forecasts.



Ref.: Gaffie, J. Abe, "Social Trends in the Soviet Union," *Amer. Journ. Soc.*, Vol. XLII, No. 3, Nov., 1936, p. 386.

S. 23 (*Continued*)

*Descriptive formula:*  $S_{23} = \underline{P}, : {}^tT^{-1} : (I)_i$

*Quantic number* = 8;1;0;1

*Legend:*

$S_{23}$  = The situation

${}^tT^{-1}$  = 10 years

records for

'| = beginning in 1925

$\underline{P}$  = the Soviet Union

$(I)_i$  = 16 indices of social trends

for each of

$$(I) = \left\{ \begin{array}{ll} {}_iIT_i^{-1} & \left\{ \begin{array}{l} \text{annual freight in } 10^5 \text{ tons} \\ \text{annual trade in } 10^7 \text{ rubles} \\ \text{population in } 10^6 \\ \text{workers in } 10^4 \end{array} \right. \\ \%P & \left\{ \begin{array}{l} \text{pupils per } 10^3 \\ \text{college pupils per } 10^3 \end{array} \right. \\ \%PT^{-1} & \left\{ \begin{array}{l} \text{annual passengers in } 10^3 \\ \text{hospital beds, rubles per } 100,000 \text{ persons} \end{array} \right. \\ IP^{-1} & \left\{ \begin{array}{l} \text{budget per capita} \\ \text{production per capita} \\ \text{coal per } 100 \text{ persons} \end{array} \right. \\ IP^{-1}T^{-1} & \left\{ \begin{array}{l} \text{steel per } 1000 \text{ persons} \\ \text{textiles, etc., per capita} \\ \text{electricity per capita} \\ \text{grain per } 100 \text{ persons} \\ \text{education, rubles per capita} \end{array} \right. \end{array} \right.$$

*Comment:*

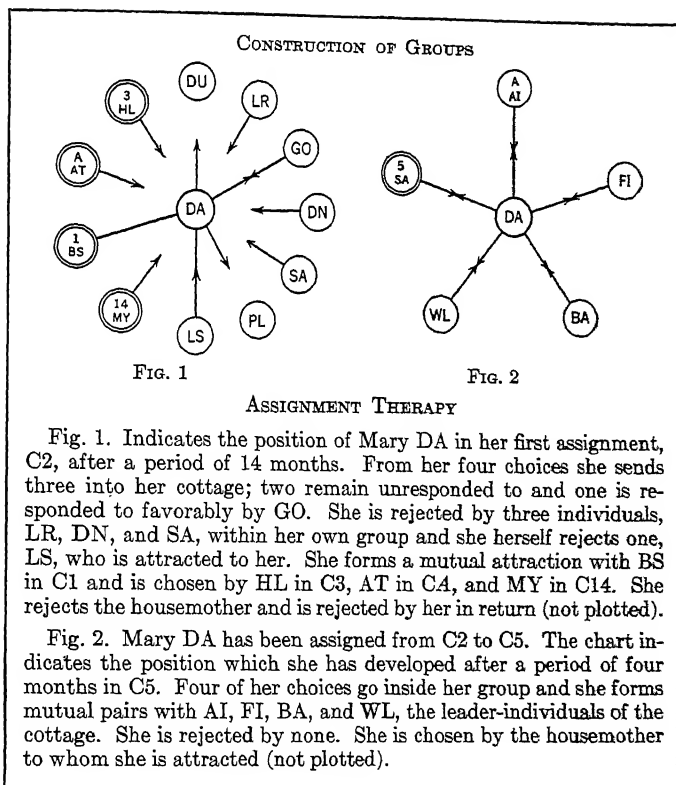
Every index trends upward except foreign trade, which is desired to go downward, so that there is progress (+,V) without exception (although grain increase is small).

The unanimity of the upward trend, even right through the depression of 1930-33, increases the probability of its continuance in the near future.

*Comment on notation:*

Note that since some of the indices are velocities, their rates of change are accelerations, making 8 the temporal quantic digit. In an aggregation the quantic digit is that of the highest exponent in the aggregated entities.

## S. 24



Ref.: Moreno, J. L., *Who Shall Survive?* Nervous and Mental Disease Publishing Co., Washington, D. C., 1936, p. 320.

Descriptive formula:  $S_{24} = {}^tT^{-1} : {}^pP :: {}^pP_{,p} : {}^iI$   
 Legend:

Quantic number = 9;1;0;2

$S_{24}$  = The situation  
 records for each of

${}^pP$  = 11 girls  
 ${}_{,p}$  = in both in-group and out-group  
 by

${}^tT^{-1}$  = 2 dates, 14 months apart  
 ${}^pP$  = one girl  
 cross-classified  
 (i.e., interrelated)  
 with

${}^iI$  = attitudes { attraction =  ${}^iI = +1$   
 in 3 de { neutrality =  ${}^iI = 0$   
 grees { rejection =  ${}^iI = -1$

Comment:

The nearer the mean attitude approaches +1 (i.e., all mutual attractions) the more complete the social adjustment of this delinquent girl in this situation,

S. 24 (*Continued*)

and the more successful towards therapy has been her assignment in the institution.

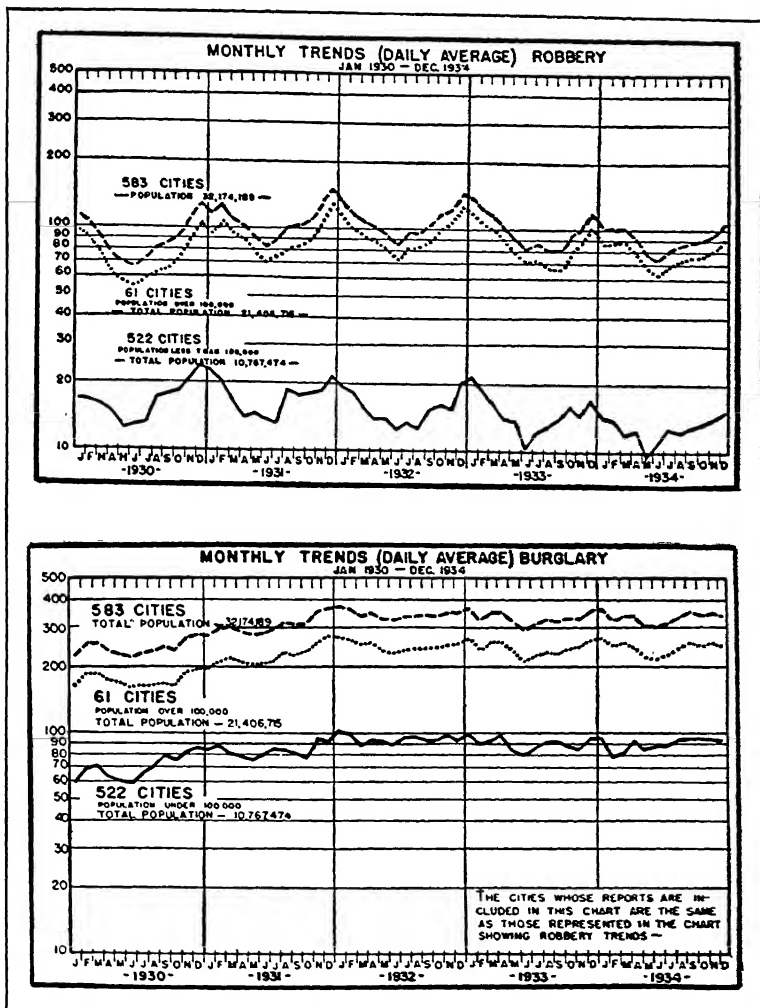
$$\begin{aligned}
 50 P^{-1} \sum_1^{2P} I &= \% \text{ of maximum adjustment} \\
 &= 50 \times 13/11 = 59\% \text{ in Fig. 1} \\
 &= 50 \times 10/5 = 100\% \text{ in Fig. 2}
 \end{aligned}$$

showing a progress of 41 percentage points (which are the I-units here) up to maximal adjustment in these fourteen months.

This illustrates a common type of prediction in the social sciences where certain actions or situations are known to result in a desired change, but just how much change cannot be predicted in advance without actuarial data.

The assignment therapy is here an instance of societal force. To estimate its amount, assume that no progress in adjustment was being made as long as this girl remained in the situation of Fig. 1, so that her velocity of progress was zero. It then increased to an average velocity of 41 I-units in 14 months, or almost 3 I-units per month. Taking the middat of this period as the truest date for this average velocity of the period, means that the velocity rose from 0 to 3 in 7 months, or  $3/7 = .43$  I-units per month per month. As the population is one person, the acceleration times 1 is the I-force. The I-force is thus estimated as .43 F-units, where the F-unit is one person improved 1 percent of maximal adjustment per month per month. But since the adjustment is an interaction between people who are explicitly specified in the situation, the force becomes societal control,  $T^{-2}IP^2$ , under our definition of control, as some people accelerating change in other people.

## S. 25



Ref.: Vold, George B., "Amount and Nature of Crime," *Amer. Journ. Soc.*, Vol. XL, No. 6, May, 1935, p. 799.

## S. 25 (Continued)

Descriptive formula:  $S_{25} = t : u T^{-1} : P_p : z_q : M(IT^{-1})_i$

Quantic number = 8:1:0:1

Legend:

$S_{25}$ = The situation	$ z_q$ = a specified number of
records for each of	cities
$t : u T^{-1}$ = 4 years and their months	$M$ = the mean
'  = beginning in 1930	$T^{-1}$ = daily
for each of	$I$ = number of police offenses
$P_p$ = 3 urban plurels	of
each of	$ _i$ = 2 kinds $\left\{ \begin{array}{l} \text{robbery} \\ \text{burglary} \end{array} \right.$

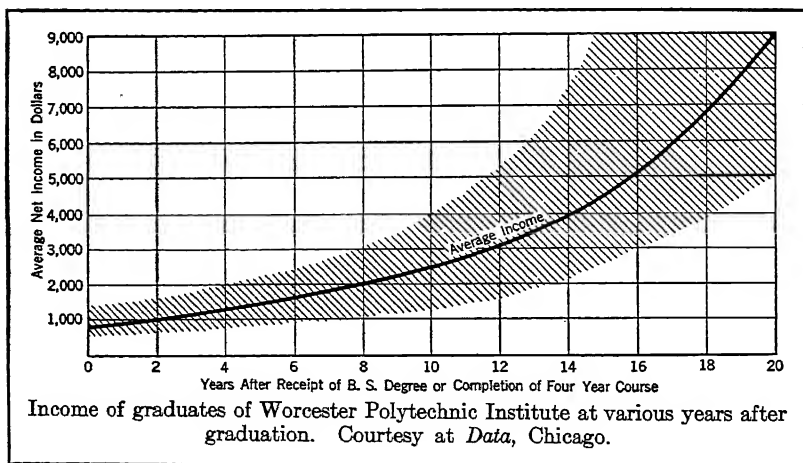
Comment:

The situation illustrates the societal force of crime of two specific types. The I-force is in F-units of one offense per urban plurel per month per month. Robberies show a seasonal cycle:

$$_t(IT^{-1} \cdot T^{+1} \neq 0) \quad (\text{Eq. 3, Ch. XI})$$

i.e., within each of the  $_t|$  years there is (curvilinear) correlation between the offenses per month and the month of the year. The correlation is high enough to justify seasonal prediction of increasing or decreasing robberies.

## S. 26



## S. 26 (Continued)

Descriptive formula:  $S_{26} = \underline{P}_t : T^{-1} : {}^{\sigma, M}(IT^{-1})$       Quantic number = 8;1;0;1  
Legend:

S<sub>26</sub> = The situation

$\bar{M}$  = the mean

records for

and

P, = graduates of Worcester Polytechnic Institute

$\sigma$  = the dispersion ( $\sigma$  is only implied by the shading)

for each of

of

 $t| = 20 \text{ years}$  $T^{-1}$  = their annual

$T^{-1}$  = the annual rate of change of

I = income

*Comment:*

This S-situation is a good example of the redispersing process. The elongating segments of ordinate in the shaded area from date to date measure the velocity of the dispersing process that is going on. Skewness is also indicated in that the shaded ordinate above the mean is longer than below the mean—the usual finding in wealth distributions.

The smooth curve represents an I-force. If fitted by an equation, the first differential of calculus would state the velocity of the earning process at any date, and the second differential would state the acceleration at any date. Since there is only one P-unit, the acceleration and the societal force are numerically equal.

## VI. NOTES

**1.** These are verbalized as:

T	denoting time	
I	“ indicators of characteristics	
L	“ length	
P	“ people	
<sup>s</sup>	the post-superscript denoting the exponent	
<sub>s</sub>	“ post-subscript	“ classes
<sub>s</sub>	“ pre-subscript	“ class-intervals
<sup>s</sup>	“ pre-superscript	“ cases
:	“ colon	“ aggregation
::	“ double colon	“ cross-classification
•	“ heavy dot	“ correlation
'	“ priming	“ particularity
+	“ plus sign	“ addition
-	“ minus “	“ subtraction
×	“ times “	“ multiplication
/	“ ratio bar	“ division

2. At the risk of repeating oneself ad nauseam, the warning as to the chief limitation of S-theory is reiterated, that these symbols represent data and do not magically convert inaccurate or inadequate data into accurately and ade-

quately represented situations. The symbols can be no better than the data they symbolize. The rules for using the symbols exert pressure on the investigator to treat his observations more precisely, but the symbols themselves passively reflect whatever the degree of precision of the data may be.

3. A comparison in respect to parsimony may be made with Eubank's four basic categories which subdivide into seven, namely (see S. 6, Ch. III) :

Societary composition—the single human being	
—the human plurel	
Societary causation	—energy (or force)
	—control
Societary change	—action (or process)
	—relationship
Societary products	—culture

The four sectors of time, space, people, and their characteristics in this volume are comparable. The scripts modifying these four and the operations combining them have no explicit concepts as equivalents in Eubank's system. He develops no explicit concepts for the *method* of modification of his basic categories into particular and variant forms and the *technic* of combining them. He presents classifications of subconcepts which do modify and do combine his basic ones without fully exposing the operational technic of such modifying and combining.

His static "Products" are the I in part of this volume, his "Composition," the P; his "Change" is reduced to T and some combining index, his "Causation" to Force, T-<sup>2</sup>IP, and "Control," T-<sup>2</sup>IP<sup>2</sup>. L is added to deal with ecology and the spatial environment more adequately. In short, the system in this volume has been built on his system, modifying its four categories into the four of S-theory which seem to be more elemental, more measurable, and more fruitful in recombining to yield current sociological concepts and descriptions of data.

4. Not all of these definitions achieve the goal of being operational definitions. In most cases it depends upon whether the indicator or the index involved is operationally derived. The S-formula usually takes that index and specifies either limiting conditions or operational combination with further indices in defining the term. Hence some of these equations describe or limit the term but cannot strictly be called definitions.

5. A fuller list is in the topical index at the end of this volume where the terms specified by formula have their equation number indicated. This list is not complete, however, as some subtypes, some phrases, and some relations between concepts are not in the topical index, yet have formulae here.

6. The S-formula for a prediction is expressed by the date script, where if 0 means by convention the present as zero point or origin, then subsequent letters in the alphabet will, as usual, denote later dates which must, therefore, be future ones.

$${}^0\text{:S} = \text{future situations, predictions} \quad (\text{Eq. 1, Ch. XII})$$

This formula asserts situations at dates, <sup>1</sup>, which come after present dates, <sup>0</sup>, to which they correspond.

7. This concept of participant prediction has much in common with Ward's teleis and other concepts of human purpose. Many of these have acquired connotations of subjective unpredictability, of "free will," of a divine Over-purpose, of vitalistic entelechy, etc. A more objective term, suitable for a behavioristic science, seems to be "participant prediction," where actuarial probabilities are connoted and mathematical treatment of the telic phenomena in human life is possible. "Planning" has much of this meaning, except that plans may be made by some parties for others—who may or may not participate fully in fulfilling the plans. Here the planners are participants, but any passive population that may be planned for are not participants. It is the active desiring and participating in achieving a social plan, not merely the making of the plan on paper (which is part of connotation of the term "social planning"), which we emphasize in the term "participant."

8. Lest the reader have the notion that repeated references to the "sixteen basic" concepts is making a fetish of this set of sixteen, it should be noted that a few more or less are acceptable. The overlining symbol, converting indices to vectors, the semicolon representing any operator, and others might be included in the set called "basic" or "chief concepts"; and minor symbols, such as the primes denoting particularity, or the double colon compounding the classification denoted by the single colon might be dropped into the "auxiliary symbols" class. The reference to sixteen basic symbols of S-theory is intended to emphasize its base in a definite number of concepts and a small number of the order of one to two dozen.

9. The probability of a prediction and its reliability should not be confused. The simple probability is the ratio of chances in 100, of an event happening. This ratio, like every statistical index, has a standard error of sampling which varies inversely with the square root of the size of the sample (Eq. 14, Ch. IX). The standard error measures the reliability with which a probability has been determined. The larger the sample the more reliable is the probability of any event computed from the experience of that sample.

10. This hypothesis that accuracy increases with relevant subclassification may be expressed crudely in the formula:

$$({}^0 : {}^1 I) \cdot {}^1 : {}^2 | {}_{\Sigma s} ? > 0$$

(Eq. 3, Ch. XII)

(See Appendix II, Rule 61)

This asserts that the accuracy index of predicting situations at particular future dates,  ${}^0 : {}^1 |$ , is positively correlated with the sum of class scripts,  $| {}_{\Sigma s}$ , corresponding to that future date. The sum of class scripts reflects both the degree and the order of the hierarchical classification, as when  $I_1 : j : k$  has three degrees of classes, so that  $| {}_{\Sigma s} = 3$ , while in  $I_1 : j$ ,  $| {}_{\Sigma s} = 2$ .  $| {}_{\Sigma s} = | {}_{\Sigma 1} + | {}_{\Sigma 2} + | {}_{\Sigma 3}$ . This assumes, of course, that the subclassifying is relevant to the predictors. It can easily be carried beyond the point of maximal accuracy of prediction and so defeat itself. Alternatively stated, the correlation may be curvilinear and reverse from positive to negative correlation at some point of optimal accuracy of prediction. The inequality, Eq. 3, Ch. XII, may hold only within limits which would have to be determined for specified situations. The six detailed

points listed under "Limits" of an index in discussing "Calibrated Indices" in the theory of measurement, section IV B of Chapter III, specify further what is involved in defining classes of the phenomena to be predicted.

**11.** Masaryk and Beneš of Czechoslovakia were outstanding exceptions.

**12.** For an excellent summary of the literature on societal control see Ref. 25. Eubank defines the term clearly (p. 219) and discusses the sources (one person on to a group), the technics (suggestion, persuasion, coercion, etc.), the instruments, the stages, and the reciprocal status of controller and controlled, and the purposes of control. These should serve as hypotheses for testing out and reformulating more exactly when more recorded situations with a quantic formula of  $T^{-2}IP^2$  are available as facts from which to induce new hypotheses and verify existing hypotheses.

# APPENDICES



## Appendix I

## GLOSSARY OF SYMBOLS OF S-THEORY

<p>16 Basic Symbols</p> <p>T I L P <math>\frac{s}{s}</math> :: : • ' + - × /</p>	<p>16 Auxiliary Symbols</p> <p>Compounds of, special cases of, or supplements to, the basic symbols</p> <p>(I) (I') S V <math>\frac{s}{s}</math>, z, M, <math>\sigma</math>; ?       ≤ ≠ ≡</p>
<p>4 INDICES</p> <p>T denotes <i>time</i>, e.g., days, years, hours, etc.</p> <p>I denotes an <i>indicator</i> of a characteristic of people, or of their environment, e.g., intelligence scores, school tests, votes, money, temperature, tons of produce, ranking in any respect, etc.</p> <p>L denotes <i>length</i> in physical space, e.g., inches, kilometers, etc.</p> <p>P denotes a human <i>population</i>, e.g., the number of persons, or plurels</p>	<p>4 INDICES</p> <p>(I) denotes any <i>index</i> defined as any combination of T, I, L, or P by addition, subtraction, multiplication, or division (but not by aggregation)</p> <p>(I') the <i>homosectoral</i> index, is any one of T, I, L, or P</p> <p>S denotes an <i>aggregation</i> of indices composing a societal situation, a set of numerical societal data (including any countable list of items, such as I<sub>q</sub>)</p> <p>V denotes a <i>valued index</i> ("desideratum"), an index of a characteristic desired by a specified population</p>
<p>4 SCRIPTS</p> <p><math> ^s =  ^{s}; i; l; p</math> = the post-superscript denotes the <i>exponents</i>, e.g.,</p> <p><math>I^3 = I \times I \times I</math></p> <p><math>I^2 = I \times I</math></p> <p><math>I^1 = I</math></p> <p><math>I^0 = I/I = 1</math></p> <p><math>I^{-1} = 1/I</math></p> <p><math>I^{-2} = 1/I \times I = 1/I^2</math></p>	<p>4 SCRIPTS</p> <p><math>^s </math> or <math>' </math> denotes the initial <i>limit</i> of a series; <math>^s </math> denotes the terminal limit of a series, e.g., <math>^s's'T</math> = bounding dates of a period; <math>^s's'I</math> = the range of an indicant</p> <p><math>^m </math> denotes the arithmetic <i>mean</i>, e.g., <math>^m(I)</math> = an average index</p> <p><math>^\sigma </math> denotes the <i>standard deviation</i>,</p>



## 4 CONVENTIONAL OPERATORS

- $+$ ,  $\Sigma$  denotes addition ( $+$  also denotes a positive quantity, an increase. The  $\Sigma$  sign before a descript means adding the list of aggregated items to get the total of such items)
- $-$  denotes subtraction, also a negative quantity, or decrease
- $\times$  denotes ordinary multiplication; a product
- $/$  denotes ordinary division; a ratio

## 4 CONVENTIONAL MATHEMATICAL SYMBOLS

- $=$  denotes "is equivalent to"
- $\equiv$  denotes "is identical to"
- $\neq$  denotes "is not equal to"
- $>$  denoting "is greater than"
- $<$  denoting "is less than"
- $||$  denotes a matrix

Other symbols which are the abbreviations of the names of societal processes and of other sociological concepts will be found in the topical index which gives references to their definitive equations and discussion in the text. For references as to the symbols above, consult the word which is underlined in the glossary above, in the topical index.

The dividing of these 32 symbols into "basic" and "auxiliary" is not rigid. One or two items might be shifted to the other side of the boundary line. They represent in general the more used and the less used, the more essential and the less essential, symbols in S-theory. Their grouping into sets of four is mostly a natural grouping as they happen to fall into classes, but is partly done as an aid to memory.

*The Meaning of the Exponent ("the Quantic") Expanded by Its Amount, Sign, and Sector*

- $|^s$  = the aggregation, or pattern, of exponents in all sectors, the quantic number ( $= |^{t,i,i,i,b}$ )
- $|^s$  = the aggregation, or pattern, of exponents (usually only a single exponent) in any one particular sector. The more conventional  $e$  (or  $|^e$ ) denotes a single exponent in one sector, whether it is in the post-superscript position or not.

$ ^s$ = the homosectoral exponent	$ ^t$ = the temporal exponent	$ ^i$ = the indicatory exponent
$ ^{s-1}$ = a variable, a quantity of a quality	$ ^{t-1}$ = a duration, an extension in time	$ ^{i-1}$ = an indicant, an amount of some characteristic
$ ^{s-2}$ = scalar product of 2 variables, including a square	$ ^{t-2}$ = a cross-classification of durations, their scalar product, their correlation	$ ^{i-2}$ = a cross-classification of two indicators, their scalar product, $\therefore$ their correlation

*The Meaning of the Exponent, etc. (Continued)*

$ s-3$ = a variable cubed $ s-0$ = a unit constant algebraically; a quality sociologically; a point geometrically	$ t-0$ = a date, a point in time if described; dateless data at any indefinite point in time if with a zero descript	$ i-3$ = a skewness index $ i-0$ = sociologically an attribute (i.e., a qualitative characteristic) undifferentiated into amounts; algebraically, a unit; geometrically, a point, one unit from zero on the indicator dimension
$ s-1$ = a ratio, a variable as divisor  $ s-2$ = a variable squared as divisor  $ s-3$ = a variable cubed as divisor	$ t-1$ = a velocity, a time rate  $ t-2$ = an acceleration, a time rate of change of a time rate	$ i-1$ = an indicant as divisor, as in an index number
$ ^1$ = the spatial exponent		$ ^p$ = the populational exponent
$ ^1-1$ = a line, a length in physical space $ ^1-2$ = an area, also the scalar product of two sets of spaces $ ^1-3$ = a volume $ ^1-0$ = a point in physical space, a dimensionless space	$ ^p-1$ = a population, a number of persons or pleurels $ ^p-2$ = the products of 2 sets of parties, the interrelation matrix, a group $ ^p-0$ = sociologically, a party (= one person or one pleurel); algebraically, a unit of population; geometrically, a point one unit from 0 on the population dimension $ ^p-1$ = a population as divisor, as in a population rate or percentage	
$ ^1-1$ = a line as divisor, "per unit of length" $ ^1-2$ = an area as divisor, "per unit of area" $ ^1-3$ = a volume as divisor, "per cubic unit"		

## Appendix II

### RULES FOR WRITING S-FORMULAE

The general procedure is first to delimit the situation for which a descriptive S-formula is to be written; then to determine what sectors are involved and determine the exponents in each sector; then to fix on the class, class-interval, and case descripts; then to write the connecting operational symbols; and finally, to check the whole formula to ensure that all rules are complied with in a consistent and clear symbolic description of the recorded societal situation in terms of the 32 symbols used in S-theory,  $(T I L P_{s;s}^{s;s} : :: \cdot ' + - \times / (I) (I') S V^a, z, M, \sigma | ; ? \_ = \equiv \neq || ||)$ . These steps are stated below in detail as rules answering the questions which the analyst should put to himself.

#### I. BASE LETTERS *S, T, I, L, P, (I)*

*Rule #1. S*—What is the situation to be analyzed?

Any convenient set of quantitatively recorded data in the social sciences may be marked off as a unit situation, *S*, for analysis. If upon analysis its quantic denotes a contiguous region in the quantic solid it may be a unit situation; if non-contiguous regions are denoted, each such region specifies a separate unit situation. A table, graph, map, or paragraph is usually the set of data marked off for trial analysis to determine the unit situation. In general, the test of a unit situation is whether it can be unambiguously described by one set of base letters and their scripts. An *S* may be a single index or an aggregation of indices.

*Rule #2. T*—Is time explicitly involved in this situation?

Write *T* if a date, period, or unit-per-time is specified (such as the date of the data, ages, income or expense per year, deaths per month, or miles per hour).

*Rule #3. I*—Is an indicator involved in this situation? (a) Every homosectoral index that is qualitatively characterized is a product of an implicit (i.e., unwritten) attribute and a pure index, and is written as that index with its scripts

$$I_1^0 P_0^{+1} \equiv P_p, I_1^0 T_0^{+1} \equiv T, I_1^0 L_0^2 \equiv L_i^2$$

*Rule #4. (b)* Write *I* if any characteristic of people or of their environment, other than time, space, population, and their implicit attributes, is asserted.

*Rule #5. L*—Is space involved in this situation? (a) Write *L* if the number of points or units of length, area, or volume are asserted, or indicated in a map.

*Rule #6.* (b) Underline as indefinite, mapped regions for which the areas are not stated numerically. (c) L is not written for areas that are merely named.

*Rule #7.* P—Is population involved in this situation?

Write P if a party or parties are named, enumerated, or canvassed, in securing the data. A party is a person or a plural.

*Rule #8.* (I)—Are different combinations of T, I, L, or P involved?

An index (I) may be written to denote a summative combination of T, I, L, or P, when such a combination is further aggregated with other indices. Addition, subtraction, multiplication, and division are included under summative combining.

Thus, a population rate is written as an index  $(PT^{-1}) = (I)$  when it is correlated with some other index in a population in order to classify that situation as a correlation,  $(I)^2$ , of two (complex) characteristics.

*Rule #9.* Write  $(I')$  and scripts  ${}_s(I')_s$  to denote a homosectoral index, i.e., an index of data from one particular sector, T, I, L, or P, for convenience in discussing them. Do not use  $(I')$  in a full descriptive formula, since the sectors that are explicit should be written out.

## II. SCRIPTS ${}_s(I')_s$

A. *Exponents* =  $|^s$  or  $e$  (self-multiplication)

$$I^3 = I \cdot I \cdot I$$

$$I^2 = I \cdot I$$

$$I^1 = I$$

$$I^0 = I/I = 1$$

$$I^{-1} = 1/I$$

$$I^{-2} = 1/I \cdot I$$

TIME—What is the temporal exponent,  $T^?$ ?

*Rule #10.*  $T^0$ . If time is stated as a single date, named instant, or specified point of time, its exponent is zero.

*Rule #11.* For timeless data also (such as " $2 \times 2 = 4$ ") the exponent is zero, but since the date script here is zero,  ${}_0$ , time is not involved in the situation as recorded, and T is not written.

*Rule #12.*  $T^{+1}$ . If a duration of time is asserted, its exponent is plus one. Ages of people at any date, and durations from variable initial dates on to a common terminal date, are denoted by  $T^{+1}$ .

*Rule #13.*  $T^{-1}$ . If an index-per-period is asserted, i.e., if an index is divided by a time duration, write  $T^{-1}$ . This denotes a speed, velocity, or time rate, of change, or of the on-going of a process. An index stated on a series of dates calls for  $T^{-1}$ .

*Rule #14.* Time rates may be for a change of static indices (which are measurable for an instant of time as a census is), or for dynamic indices, such as events, acts, income, and expenditure, which tend to vary proportionately with the length of the period observed and vanish if only a point of time is considered.

*Rule #15.* Durations from a common initial date on to either a common date, or to variable terminal dates are considered as emphasizing change,  $T^{-1}$ , rather than duration,  $T^{-1}$ .

*Rule #16.*  $T^{-2}$ , (or  $T^{-1}T^{-1}$ ). If an index-per-period changes from one period to another, write  $T^{-1}T^{-1}$ . This denotes a celeration. To distinguish velocity from celeration, always consider the data as plotted with time as one axis (abscissa). Then if the ordinates are static indices, the situation involves a velocity; if the ordinates represent dynamic indices, or any velocity, the situation involves acceleration.

*Rule #17.*  $T^{-1}T^{-1}$ . If both change and duration are involved, write  $T^{-1}$  and  $T^{-1}$  separately, each with appropriate descripts.

#### INDICATORS—What is the indicatory exponent, $I^?$

*Rule #18.* For a qualitative characteristic, an attribute, the exponent is zero,  $I^0$ . Every characterized index is an implicit product of  $I^0$  and the pure index, but the  $I^0$  is considered implicit (i.e., understood) and is not written. Qualitative characteristics or attributes other than those naming the index are written explicitly. Examples: negro working population =  $P$ , (really  $I^0I^0P_0$ ); types of machines  $i$  in number, with number of workers for each type  $I_i^0 : P$ , (really  $I_i^0 : I^0P_0$ , where  $I^0$  = "working" and  $P_0$  = "persons").

*Rule #19.* For a quantitative characteristic (indicant) the exponent is one,  $I$ , and may be understood without writing it. Quantitative characteristics include all-or-none attributes, ordinals (ranks, or unequal degrees), and cardinals (multiples of equal units).

*Rule #20.* For an index number, i.e., an indicant divided by an indicant, the quantic digit is plus one.

*Rule #21.* A division merely to alter the units to relative terms within one sector is not considered in determining the quantic digit, i.e., the exponent.

*Rule #22.* For two characteristics whose correlation is stated in scatter, or coefficient, or other form, the exponent is 2. This is written in the descriptive formula by repeating the index  $(I)^{+1}(I)^{+1}$ . For the variance, and for any second moment, the exponent is also 2.

*Rule #23.* For a third moment  $e = 3$ , for the  $n$ -th moment  $e = n$ .

#### SPACE—What is the spatial exponent, $L^?$

*Rule #24.* For a point the exponent is 0.

For the *number* of spaces the second exponent is zero and the class script is summed. Thus,  $(L^2)_{\Sigma}^0$  = areas, taken as units, 1 in number.  $_{\Sigma 1}(L^3)^0$  = the number 1 of equal volumes. In case of a second exponent both are written in the quantic number—thus  $|^s = 0;0;20;0$  and  $0;0;30;0$  for the examples above.

*Rule #25.* For a line the exponent is 1.

For a mean of a series of spaces the second exponent is 1. Thus,  $\Sigma(L^2)/_{\Sigma 1} = {}^ML^2 \quad |^s = 0;0;21;0$ .

*Rule #26.* For an area the exponent is 2.

Also for a second moment of spaces (i.e., variance or correlation of lines, areas, or volumes) the second exponent is 2.  $\Sigma(L^3)^2/\Sigma_1 = (\sigma L_1^3)^2$   
|<sup>s</sup> = 0;0;32;0.

*Rule #27.* (a) For a volume the exponent is 3. The exponent of 3 also denotes the third moment of a distribution of spaces.

(b) The quantic digit denoting the physical dimension is written first and the digit denoting the statistical moment is written second, since this is the sequence of performing the exponential operations.

*Rule #28.* For any index divided by units of length, area, or volume the exponent is -1, -2, and -3 respectively.

POPULATION—What is the populational exponent, P<sup>p</sup>?

*Rule #29.* If no population is named or enumerated, the exponent is zero. (And the class script is zero denoting a "nul" population, i.e., one that is non-existent in that situation, P<sub>0</sub><sup>0</sup>, and is not written in the descriptive formula.)

*Rule #30.* If a population is named, or enumerated, or canvassed in obtaining the data, the exponent is 1.

*Rule #31.* If the ratio of two different populations (other than a part to the whole population) is stated, the exponent is plus and minus one, P<sup>+1-1</sup>. The quantic number records both exponents, thus |<sup>s</sup> = 0;0;0;19. For a part of a population expressed as a ratio to the total, this shift from absolute to relative units is not considered as altering the quantic digit of 1. Cf. Rule #21.

*Rule #32.* If *stimulus-response* interrelations of parties with parties are asserted, the populational exponent is plus 2 in the quantic formula and is written in the descriptive formula as a repetition of the P, i.e., as P :: P.

QUANTIC—What is the quantic number?

*Rule #33.* The four exponents |<sup>q:1:1:P</sup>, symbolized by |<sup>s</sup> are called the quantic number of a situation.

- (a) A "nullary" index is one with an exponent of zero
- (b) A "primary" " " " " " " " " 1 or 9
- (c) A "secondary" " " " " " " " " 2 or 8
- (d) A "tertiary" " " " " " " " " 3 or 7

*Rule #34.*

For e = -1 write the digit 9

For e = -2 or -1-1, write the digit 8

For e = -3 write the digit 7

*Rule #35.* If more than one exponent occurs in one sector, (a) with cross-classified or correlated bases, add the exponents to obtain the quantic number, e.g.,

$$I, \bullet I,, \quad |^s = 0;2;0;0, \text{ and } P_p :: P_p \quad |^s = 0;0;0;2$$

*Rule #36.* If more than one exponent occurs in one sector, (b) with summated or aggregated bases, write all the exponents to obtain the quantic number. Examples:

1. positive and negative exponents

$$P_i^{+1}P_{j-1}^{-1} \quad |^s = 0;0;0;19$$

2. a second operation of exponents, as in computing a statistical moment of spaces

$$\Sigma(L^{+1})^2/|_1 \quad |^s = 0;0;12;0$$

3. an aggregation of spaces, as areas and lines

$$L_i^2 : L_j^{+1} \quad |^s = 0;0;21;0$$

4. contact of interrelated parties

$$^pP : ^qP \quad |^s = 0;0;0;11$$

5. a scattergram for a correlation ratio

$$_iI : ^MI \quad |^s = 0;11;0;0^*$$

*Rule #37.* A situation is classified in the quantic classification by the first digit in each sector of the quantic number.<sup>†</sup> Hence in a sector the dominant exponent is written first in the quantic number. See examples under Rule 36.

### B. Descripts (Aggregations) $\frac{s}{s}|_s$

*Rule #38.* For more than one aggregation requiring a classification or hierarchy of descripts, use the sectoral letter for the first aggregation and the letters following it in the alphabet for the next aggregations, e.g.,

$$I_i : j : k, L_l : m, P_p : q : r : s, t : u : vT$$

*Rule #39.* Let the descripts  $(\frac{s}{s}|_s)$  symbolize the pattern of descripts in the whole situation, and let this together with the prime denote the pattern of scripts in any one sector  $\frac{s}{s}|\frac{s}{s}$ .

CLASS SCRIPTS  $|_s$ —What is the number of classes, i.e., qualitatively different indices, in each sector?

*Rule #40.* Write  $|_t$  to denote the number and kind of chronologies, or sets of time units.

*Rule #41.* Write  $|_i$  to denote the number and kind of aggregated indicators, i.e., recorded characteristics of people, or of their environment other than T, L, and P.

\* It may prove desirable to adopt some convention or rule to symbolize differentially a second exponent when the bases are aggregated, added, multiplied, or raised to a power. At the present stage of development of S-theory this refinement does not seem necessary. The quantic digits by this rule, as stated above, reflect the structure of the situation with very little ambiguity.

<sup>†</sup> Should situations, not yet encountered, occur with exponents in such a way as not to fit into the categories above, or as to leave the analyst doubtful as to the correct quantic, the digit 5 may be arbitrarily reserved for denoting such, in the quantic number. (5 might class with 1 in that it usually denotes some complex index.)

*Rule #42.* Write  $|_1$  to denote the number and kind of aggregated spatial regions.

*Rule #43.* Write  $|_p$  to denote the number and kind of aggregated qualitatively different plurels,  $p$  in number, e.g., a list of nations, a series of occupational plurels, etc.

*Rule #44.* For regional plurels recorded by a map, the qualitative characteristics which identify them are denoted by the  $|_1$  script.

*Rule #45.* Whenever desired, other letters defined by the legend may be substituted for an index with a particular class script. Thus, for a valued indicant  $V$  for  $L$ ;  $X$ ,  $Y$ ,  $Z$ , or  $I$ ,  $J$ ,  $K$ , for  $I$ ,  $I_{,,}$ ,  $I_{,,,}$ , etc.

CLASS-INTERVAL SCRIPTS,  $|_s$ —What is the number of class-intervals?

*Rule #46.* Write  $|_T$  for a series of periods of time,  $|_t$  in number, each of duration  $T$ , e.g.,

$|_T^{T+1} : P$  a population age distribution  
 $|_T^{-1} : (I)$  a time curve of an index  
 $|_t : |_u T^{-1}$  a series of periods and subperiods

*Rule #47.* Write  $|_I$  for a series of class-intervals,  $|_i$  in number, of an indicant. This denotes the aggregation of class-intervals, i.e., the scale itself. The  $I$ , when modified by the pre-subscript,  $|_i$ , denotes the number of  $I$ -units in each class-interval. This descript cannot occur on  $I^0$ , as it denotes quantitatively equal class-intervals, e.g.,  $|_i I$  = wealth in 5 class-intervals of \$1000 each, representing a series from 0 to \$5000; or again,  $|_i I$  = school grades in 20 class-intervals of 5 percentage points each, representing a distribution ranging from 0 to 100%.

*Rule #48.* Write  $|_1$  for an aggregation of equal sects ( $|_1 L$ ), or equal units of area ( $|_1 L^2$ ), or equal units of volume ( $|_1 L^3$ ).

*Rule #49.* Write  $|_P$  for the series of class-intervals of population whenever quantitatively equal and qualitatively interchangeable plurels are the units. Thus, for an aggregation of incomes by families,  $|_P P : I$ ; or for the number of persons in each income class-interval,  $|_I : P$ .

*Rule #50.* Special units may be specified by special symbols, thus :

$|_s I$  = an indicant in standard deviation units  
 $|_s I$  = an indicant in dollar units  
 $|_P I$ ,  $|_P P$  = relative units such as a percent, index numbers, persons per 1000, which are all percents or further decimals of percents. (The quantic digit is 1.)  
 $|_p |$  = plurels as units, i.e., families, schools, churches, corporations  
 $|_? (I)$  = an index in hypothetical units, i.e., uncertain or even unknown as yet

CASE SCRIPTS,  $|_s$ —What is the number and kind of aggregated specified cases, or points, on the index?

*Rule #51.* Write  ${}^tT$  for the number of specified dates, or for the number of punctiform events in sequence where their time intervals may or may not be equal.

*Rule #52.* If the dates are an aggregation not in sequence, write  ${}^tT$ . See S. 45, Ch. X. The quantic digit may be 0, 9, or 1 depending on Rules 10-17.

*Rule #53.* For a series of consecutive periods of unequal length, write  ${}^a{}_tT$  to show that the initial dates of the varying periods are stated.

*Rule #54.* For a series of non-consecutive periods, equal or unequal, write  ${}^{a:z}{}_tT$  to show that the initial and terminal dates corresponding to the periods are specified.

*Rule #55.* Write  ${}^iI$  for the number of listed points such as mean, sigma, or other points on an indicant. If the units of the indicant are ordinal units such as ranks, or an all-or-none attribute, they should be symbolized by the case scripts instead of the class-interval script, as the latter denotes cardinal units, i.e., equal standardized units, e.g.,

${}^iI$  = a series of ranks in some characteristic  
 ${}^2I$  or  ${}^1{}_0I$  = an all-or-none characteristic

*Rule #56.* Write  ${}^lL$  for the number of aggregated points in a line ( ${}^1L^1$ ), area ( ${}^1L^2$ ), or volume ( ${}^1L^3$ ).

*Rule #57.* Write  ${}^pP$  for the number of persons in an aggregation. This denotes a list, i.e., aggregation of specified persons, while  $P$  denotes a single number of persons.

*Rule #58.* Write  ${}^{a:z}{}_t$  for the initial and terminal limits. For the singular case,  ${}^{a:z}{}_t$  may be condensed to  $' : z'$ , e.g.,

$' : z'{}_tT^{+1}$  = 10 5-year age periods from age 20 to age 70

$' : z'{}_tI$  = scores on an intelligence test in the range from 40 to 160 score-points, etc.

### III. OPERATORS

What is the form of combination, i.e., what are the mathematical operations for combining the indices and scripts?

*Rule #59.* Use the conventional symbols with their usual meaning in mathematics:

equality, or equivalence, =  
 inequality,  $\neq$   
 "greater than,"  $>$   
 "less than,"  $<$   
 identity,  $\equiv$   
 addition,  $\Sigma$  or  $+$   
 subtraction,  $-$   
 multiplication,  $\times$   
 division,  $/$  or  $|^{-1}$   
 unitary treatment,  $( )$

a matrix  $|| \quad ||$   
 an absolute value regardless of algebraic sign,  $| \quad |$   
 separation of items, a comma ,

*Rule #60.* (cf. Rule 39) (a) The summation sign in a descript denotes the sum or the number of the entities aggregated.

Thus,  $\sum_s' \sum_s = \sum_{i+j} \sum_{i+j+k} \text{ or } \sum_{p+q+r} \sum_{p+q} \text{ or } L_{1+m+n}^2, \text{ etc.}$

(b) This descript,  $\sum_s$ , can state the order of the matrix, as far as determined by the index to which it is attached. Thus, in frequency distributions of 5 characteristics each in 20 class-intervals and observed in each of 4 plurels, the descriptive formula is  ${}_1I_1 : P_p$ , and the order of the matrix is :

$$\sum_i | \times | \sum_i \times | \sum_p \text{ or } 20 \times 5 \times 4 = 400.$$

(c) The order of the matrix *for the class scripts* is ordinarily the number of dimensions in the situation, i.e.,

$$|\sum_i + |\sum_p = 5 + 4 = 9 \text{ in this illustration}$$

*Rule #61.* The summation sign with the capital letter S in a descript may be used to denote the degree of the matrix, i.e., the number of letters in descripts:

$$\sum_s' \sum_s \text{ for any one sector, and } \sum_s \sum_s \sum_s \text{ for the whole situation}$$

Thus,  $\sum_s | = 2 \text{ in } p : q |$   
 $\sum_s | = 3 \text{ in } i : u : v | \text{ and}$   
 $|\sum_s = 4 \text{ in } | p : q : r : s$

*Rule #62.* The colon, :, denotes aggregation, meaning that each of the aggregated items preceding the colon has corresponding further aggregated items as denoted by the symbols following the colon, e.g.,

${}^tT^{-1} : P =$  a population at each of  $t$  dates

${}^tT^{-1} : P_p =$  an aggregation of  $p$  plurels each with a specified number of persons,  $P$ , on each of the  $t$  dates

*Rule #63.* The index preceding the colon is the independent or dominant variable for which dependent values are observed and recorded by the index following the colon, e.g.,

${}_1I : P =$  a frequency of persons corresponding to (i.e., depending on) each of the class-intervals of the indicant

$P_p : I =$  for each of  $p$  plurels (each of which comprises a stated number of persons) there is an amount of an indicant

*Rule #64.* The colon denotes subclassification, especially in the descripts, e.g.,

$P_p : q =$   $p$  listed plurels, each subclassified into  $q$  others

*Rule #65.* The double colon, ::, denotes cross-classification of two aggregative indices, as in a correlation scattergram,  ${}_1I :: {}_jI : P$ , or an interrelation matrix,  ${}^pP :: {}^pP : I$ .

*Rule #66.* A subclassification of one set of classes into a second set becomes a cross-classification when it is reversible, i.e., when the data are so presented that the first set of classes could equally well be considered as subclasses of the second set of classes.

*Rule #67.* Within a sector subclassification is denoted by the colon in a script, but for cross-classification the base letter is repeated with a double colon between them.

*Rule #68.* Cross-classification steps the exponent (the quantic digit) up one integer. (Thus, for  $I_i : j \mid_s = t; 1; j; p$ , but for  $iI :: jI \mid^s = t; 2; j; p$ .)

**NUMBER**—Is the number of indices in a sector singular or plural, one or more?

*Rule #69.* For one particular index add, or substitute, a prime to the descript.

*Rule #70.* For the second particular (i.e., identified) index add, or substitute, a double prime; then in general:

- $iI$  or  $,I$  = a particular class-interval
- $P_{p''}$  or  $P_{''}$  = a second particular plurel
- ${}^{p'''}P$  or  $'''P$  = a third particular person
- ${}^{iv}T$  or  ${}^{iv}T$  = a fourth particular period
- $I_{iv}$  or  $I_v$  = a fifth particular indicant, etc.

*Rule #71.* For an aggregation of indices in a sector write a small letter as descript. An aggregation is a series, a list, an itemized collection of numbers, which are not added but are merely tabulated, or collected.

*Rule #72.* For a single number denoting the number of indices (or class-intervals or cases) use a summation sign,  $\Sigma$ , before the descript, i.e.,

- $I_i^0$  = an aggregation of attributes of  $i$  kinds
- $I_{\Sigma i}^0$  = a number,  $i$ , of attributes

*Rule #73.* In S-theory a small letter descript always denotes an aggregation, a list of entities, a collection of numbers; a summation sign before a small letter in the script denotes an amount, a single number; a prime on a small letter denotes a particular identified instance of that aggregation, e.g.,

- $P_p$  = a list of plurels,  $p$  in number
- $|_{\Sigma p}$  = a single number, the number of plurels
- $P_p$  or  $P_p'$  = one particular plurel in the list
- $|_{\Sigma i}$  = the number of different indicators
- $I_{\Sigma i}^{+i}$  = the sum of the values of the  $i$  different indicants

*Rule #74.* The dot,  $\bullet$ , denotes a scalar product of two vectors and statistically denotes the covariance (which when in sigma units becomes the correlation coefficient). Hence the dot also denotes correlation.  $\bullet I \bullet J \equiv r_{IJ}$ . This may also be written as an index which is the  $r$  between a first and a second particular index ( $I$ ),  $\bullet$ ,  $''$ .

actors and the recipients respectively, or the first particular factor and the second particular factor respectively in the quantic product,  $P^2$ . The quantic is  $I^{+1}P^{+2}$  for this and for the case where plurals replace persons in the interrelation matrix, i.e.,  $P_p :: P_q : (I)$ ; and  $P_p :: P_P : (I)$  (or  $I_p :: q$  and  $I_p :: P$  in Brief-S formulae).

Rule #98.  $tT^{-1} : I_1$  = time curves  $|_i$  in number, compared. As correlation is only indirectly involved through the time component, the indicants are written as aggregated, not as correlated ( $= I_1$  in Brief-S).

Rule #99.  $I^0T^{-1}$  = an event, a qualitative action (a verb).

Rule #100.  $L_1^2 : L_1$  = a map of areas with specified lines on it.  $|^s = 0;0;21;0$ . See Rule #36.

These rules for writing S-formulae may be illustrated by detailed analysis of some sample situations as follows:

Consider S. 1, Ch. II. By Rule #1 the two circles and captions of "Fig. 4" and "Fig. 5" are marked off as the "quantitatively recorded situation" to be described by an S-formula. By Rule #2 a T is written. By Rule #3 the qualitative characterizing of time into *kinds* of ages is expressed as the product of an implicit attribute and pure time, and is written as T with qualifying scripts. The references to space are merely a name "the earth" and are, therefore, not written by Rule #3. The population reference also seems too remote from an enumeration or canvass to justify writing it explicitly by Rule #4.

Proceeding to consider exponents, the time is a duration, and by Rule #12 is so denoted by an exponent of plus one,  $T^{+1}$ . Considering descripts next, two unequal periods are found, one of which is subdivided into 2 further unequal periods. The period script called for by Rule #46, is subdivided by using the next letter of the alphabet by Rule #38, yielding  $tuT^{+1}$  thus far. Since the periods are unequal, this is denoted by asserting a series of initial dates by writing  $tu^aT^{+1}$  by Rule #53. Finally, since the periods are subclassified, the colon relates them, yielding the final formula  $S = t : tu^aT^{+1}$ , for which the quantic number by Rules #33 and #34 is  $1;0;0;0$ .

For a second example consider S. 15, Ch. II, census data punched on the older type of card for a tabulating machine. "Fig. 228" is the S by Rule #1. Time is involved as a date and so is written as  $T^0$  by Rules #2 and #10. A party is involved identified by the number "1263," so  $P^{+1}$  is written by Rules #7 and #30. The entries are an aggregation of characteristics, some qualitative, some quantitative, so that  $I^0$  and  $I^{+1}$  are both involved by Rules #4, #18, and #19. Space is not written by Rule #6. We now have  $T^0I^{+1}P^{+1}$  and proceed to consider the descripts. For the aggregation of characteristics, write  $I_1^0$  by Rule #41. By Rule #62 writing the colon and  $I^{+1}$  right after  $I_1^0$  would mean that some of these attributes have been quantified, i.e., have corresponding amounts of their quality.  $I^{0,1}$  is a condensation for  $I^0, I^{+1}$ , an aggregation of qualitative and quantitative indicators. Since the date and the person are specified, a

prime is written by Rule #69 in the date script on the T and in the person script on the P, yielding  $T^0 I_1^0 : I^{+1} P^{-1}$  thus far. Since to the one date there corresponds one person with his corresponding characteristics, the colon connects the items of this aggregation by Rule #62. As the characteristics are dependent on the person, the P precedes the colon, while the I's follow it by Rule #63, and all follow the date relative to which the data are determined. This gives finally:  $S = T^0 : P : I_1^0$ , with a quantic number by Rule #33 of  $0;1;0;1$ .

For a third worked-out example of applying the rules for describing a situation in standardized symbols, consider S. 18, Ch. II. presenting the death rates in various forms of travel. The paragraph from "You may . . . accident victims" on page 271 of Furnas' book is defined as the S-situation for analysis by Rule #1. Time, length, population, and characteristics are involved by Rules #2, #3, #5, and #7. Time and length are in the denominator, as the data are in units per year per mile, and hence have exponents of  $-1$  by Rules #13 and #28. Population is a ratio of persons dying to total persons and may be written as  ${}_cP$  by Rule #50. The qualitative characteristics of five types of transportation,  $I_1^0$ , may be implicitly written as five types of transportation plurals by the class script on P by Rule #3. The combination of the TL and P indices is by multiplication into the heterosectoral index  $(T^{-1}L^{-1}{}_cP)$  by Rules #59 and #85. The order is immaterial in an arithmetic product. The aggregation of the five indices for the five plurals gives by Rule #43 the final formula as  $S = ({}_cPL^{-1}T^{-1})_P$  with a quantic number by Rules #33, #34, and #31 of  $9;0;9;1$ , classifying the situation as a lineate de-populating process.

## *Appendix III*

### RESEARCH SUGGESTIONS

The following list of unfinished tasks, of hypotheses to be tested, and of fields to be explored, may be of use to the student who wants to pursue further original investigation in the field of this volume.

#### *I. GENERAL PROBLEMS*

1. Most of the formulae in this volume are hypotheses which need to be tested. With the exception of the formulae here which are found in statistical textbooks and a few other mathematically derived formulae, most of the equations propose symbolization for a certain societal content. By applying these formulae to large collections of appropriate data, their degree of adequacy and criteria for this adequacy may be sought. The suitability of the verbal labels naming the formulae may also be scrutinized. Extensive subclassification and intercorrelating of the formulae should be further results of these investigations.

2. A mathematical textbook for sociologists is needed. At present the diverse branches of mathematics used in such a study as S-theory require a major in mathematics to master. But selections from Algebra and Geometry (plane, solid, co-ordinate and n-dimensional), Trigonometry, Calculus, Matrix and Vector Algebra, and Statistics including component analysis and actuarial formulae, could be put together into a course of mathematics for the social sciences. A cycle to develop general understanding of principles and a further cycle to develop detailed computational skill in applying these to data might be distinguished in such a textbook.

#### *II. PROBLEMS CONNECTED WITH CHAPTER II— THE S-THEORY*

##### *A. Systematization*

3. Systems based on other than the four sectors used here might be developed and tried out. Three sectors treating  $L_1^1$  as  $(I)_1$ , or five sectors including money, M, etc., might be explored in a thousand situations to see what advantages or disadvantages emerged.

4. Comparisons with other conceptual and symbolic systems may be made.

a. Thus, Eubank's and Von Wiese's systems might be translated into S-symbols with emerging similarities and differences noted.

b. The extent of overlap between S-theory's societal categories and

those of a general philosophic system such as Kant's categories of quantity, quality, relation, and modality might be explored.\* The superiority of one or the other might be experimentally tested by the percent of agreement of two independent investigators classifying 1000 suitable items into these sets of categories.

c. The integration of S-notation and some system of symbolic logic might be attempted. See the hypotheses on this in relation to Whitehead and Russell's *Principia Mathematica* in a footnote in Chapter III.

5. The rarer exponents of the quantic system might be rigorously explored. What is the possible meaning of exponents of +3, +4, +5, -2, -3, fractional exponents, etc., in each sector? To what extent have situations been published corresponding to these unusual quantics?

### B. Utility

6. A popularizing textbook testing whether the notation and classification of S-theory can be used fruitfully by beginning college students in Sociology would be one test of the utility of the theory.

7. The thoroughgoing application of S-formulae might be attempted for each conventional unit subfield of Sociology, clustered around a topic

\* A more explicit lead can be secured from the following tabulation:

<i>Kant's Categories</i>		Possible or partial equivalents in S-theory to be investigated as far as societal phenomena are involved
<i>p. 107 †</i>	<i>p. 113 †</i>	
<i>Quality</i>	<i>Of Quality</i>	$I^0$ = a qualitative index
Affirmative	Reality	$I^0 = 1$ , something affirmed
Negative	Negation	$I^{-\infty} = 0$ , something negated
Infinite	Limitation	$I^{+\infty} = \infty$ , something infinite
<i>Quantity</i>	<i>Of Quantity</i>	$(I)^{+1}$ = a quantitative index
Singular	Unity	$'(I)$ = one case
Particular	Plurality	$\wedge(I)$ = one class-interval, or part of all
Universal	Totality	$(I)$ , = one class, the whole of one kind
<i>Relation</i>	<i>Of Relation</i>	$(I)^2$ = a relative index
Categorical	Of inheritance and subsistence	$I^2$ = correlation of characteristics
Hypothetical	Of causality and dependence	$T^{-2}$ = acceleration, which with $I^2$ = causation
Disjunctive	Of community (reciprocity between agent and patient)	$P^2$ = interrelations of persons in communities and other groups
<i>Modality</i>	<i>Of Modality</i>	$(I)^i$ = special subforms of indices
Problematic	Possibility-impossibility	$I^{+1-1}$ = a probability ratio
Assertoric	Existence-non-existence	$I^{0,-\infty}$ = something existent, something non-existent
Apodeictic	Necessity-contingency	$I^2$ = a correlation of 1.0 or less

† Smith, Norman Kemp, *Immanuel Kant's Critique of Pure Reason*, Macmillan, 1929, page 681.

such as ecology, rural Sociology, population larithmics, culture, community, social organization and disorganization, social lag and strain, institutions, etc. Can all the quantifiable data in each subfield be expressed in S-formulae? Does improved classification, objectivity, precision or fruitfulness in stating hypotheses, or in exposing relations, or in determining exact degrees of relationship, or in exposing inadequacy of the data, emerge?

Similar studies for cognate social sciences—Anthropology, Economics, Political Science, Social Psychology—might be made.

8. Can dimensional analysis of societal situations be used, as dimensional analysis is used in Physics (see Bridgeman, P. W., *Dimensional Analysis*, Yale University Press, revised ed. 1931, p. 113), to check upon the soundness and the internal consistency of our analyses and comparisons and manipulations of societal situations?

9. What further concepts (see lists in Chapter XII) of use in the social sciences can be defined or be given increased specification by combinations of the symbols of S-theory? If other symbols were to be added, would the range of concepts then symbolizable be proportionately enlarged?

### C. Verification

10. Reliability. What is the percentage of agreement between other pairs of analysts on analyzing 1000 S-situations into their descriptive and quantic formulae under conditions of training varying from merely studying this volume up to adequate background in statistics and the social sciences, plus a course in S-theory involving supervised practice in S-analysis?

11. Geometrical consistency. Can tensor theory be applied to S-theory? Can vector products of vectors be utilized? Can some non-Euclidean geometry deal more adequately with qualitative phenomena ( $I^0$ ); with hierarchies of dimensions,  $|_s$ ; with the synthesis of exponents from the statistical moments and from physical space ( $L^1$ )<sup>1</sup>; with negative exponents; with aggregates of different exponents in one sector, etc.?

12. Validity. A published reservoir of several thousand S-situations, reliably analyzed into their formulae, would fulfill several functions such as:

- a. to test further the correspondence of the theory to societal facts;
- b. to provide teaching material for students studying S-analysis;
- c. to standardize S-notation and the rules of notation further;
- d. to provide data for further research by providing collections of situations defining societal "control," "community," etc., from which further inductions can be made.

13. A questionnaire to students and critics of S-theory later on should reveal what its weak or controversial or unexplored points are and which, if any, of these merit further research and in what order of priority. This should also reveal research published or in progress, indicating the more

fruitful parts of the theory or of the field to which the theory is intended to contribute.

### III. PROBLEMS CONNECTED WITH CHAPTER III, INDICATORS

14. The attribute hypothesis,  $I^0 = \text{qualities}$ .

a. What are the limits of this definitional assumption? Should it denote anything namable by a word, or should it be reserved for characteristics for which indicators of some specified reliability and validity have been developed?

b. What are the best operational rules to standardize for combining (i.e., adding, subtracting, multiplying, dividing, aggregating, cross-classifying, correlating, equating, etc.) attributes with each other and with quantitative indices?

15. Our elaboration of Bernard's classification of indicators according to content needs refining towards the goal of becoming an orderly presentation of all the raw data with which a sociologist may be concerned in the form of calibrated indices, as fully as the current state of science may have developed such calibration.

16. The increasing-precision-of-measurement hypothesis (that observation proceeds from attributes through ordinals to calibrated cardinal indices) needs fuller testing. Are there variant observations not included in this series? The boundaries between members of the series need sharpening by controlled experiments. Specific technics for calibrating instruments for societal observation need to be invented or further amplified.

17. The fundamental research problem, larger in magnitude than all the others listed here when put together, is to develop technics for securing operationally defined societal data. This is the endless and multiform process of inventing and organizing so that attributes can be observed as ordinal indicants and these as cardinal indicants, and these as further calibrated ones. This process of observing societal facts more accurately and in a form which is less dependent upon words-naming-opinions is the basic process whereby every science makes progress. It means the laborious and competent devising of schedules and scales to define eventually every concept used in Sociology. Here is work for generations of graduate students in hundreds of colleges and research agencies.

### IV. PROBLEMS CONNECTED WITH CHAPTER IV, PLURELS

18. Criteria are needed for sharper differentiating of plurels,  $P_p$ , from groups,  $P_p^2$ , when interrelations are inferrable but are not reported in the data as recorded, and for better standardized differentiating of explicit plurels,  $P_p (= I_1^0 P_0^{+1})$  from nul plurels,  $I_1^0 P_0^0 (= I_1^0)$ .

*V. PROBLEMS CONNECTED WITH CHAPTER V,  
DISTRIBUTIONS*

19. The hypothesis of a natural range of  $12.5\sigma$  should be explored on biometric, psychometric, and sociometric data.

20. The properties of epsilon elements which yield distributions which are not normal (Gaussian) might be worth discovering.

21. Following Pemberton's lead, normal ogives might be fitted to other cultural data. Wherever, as in immigration in S. 3, Ch. XI, the situation is unfinished, prediction may become possible.

22. The tension theory should be explored. Will a difference form of equation ( $PD - V = E$ ) handle unlimited, negative, qualitative, and other special cases better than the ratio form? What correlations exist between the factors under specified conditions, limiting the number of degrees of freedom? Can this tension theory be related to any one or all of the economic value theories of cost of production, marginal utility, etc.?

*VI. PROBLEMS CONNECTED WITH CHAPTER VI,  
CORRELATIONS*

23. The epsilon ( $\epsilon$ ) elements hypothesis should be explored. Can experimental or statistical conditions be devised whereby the numbers of such elements or societal atoms, constituting observed indices and their components, may become determinate?

24. Stevenson's "Q-technic" (S. 24, Ch. VI) might be explored.

25. A large number of problems in component analysis might be suggested, but there are already many active workers in this field with a special journal, *Psychometrika*, as their organ.

*VII. PROBLEMS CONNECTED WITH CHAPTER VII,  
INTERRELATIONS*

26. The technics of matrix and vectorial algebra might be applied to interrelation matrices. What is the equivalent here of a scalar product, the correlation coefficient of indicant matrices? Will component analysis yield significant results?

27. The equations for types of leadership and for the boundaries between subclasses, or standardized degrees, of isolation, contact, and for different particular indices of relationship might be worked out.

*VIII. PROBLEMS CONNECTED WITH CHAPTER VIII,  
DENSITIES*

28. The equation roughly relating density with type of economic productive culture might be utilized in making a trial of the technic of converting an ordinal indicant into a calibrated cardinal indicant by means of a precise and operationally defined relation to a second indi-

cant (Dn). Compare the operational definition of the Centigrade scale of temperature where the behavior of one variable (mercury) is calibrated by its relation to the behavior of another variable (water, when it freezes or boils).

#### IX. PROBLEMS CONNECTED WITH CHAPTER IX, DURATIONS

29. The technics of insurance actuaries might be translated into S-formulae to determine the extent to which sociological theory and practical applications in societal predicting can be integrated.

#### X. PROBLEMS CONNECTED WITH CHAPTER X, CHANGE

30. All the processes as defined here are hypotheses which, as noted in #1 above, need to be explored.

31. Sorokin's mobility propositions might be more rigorously tested.

32. Tarde's imitation hypotheses might be more rigorously tested, now that more exact formulation has been achieved.

33. The tension hypothesis (Eq. 39a, Ch. X) of the tension theory needs exploring to determine the degree of correlation between societal tension and action under specified conditions. (Cf. #22 above.)

34. Situations involving competing, Cp, should be collected and studied in order to induce further sociological generalizations. The theory on co-operative and competitive habits developed by the Social Science Research Council's committee (May, Mark A. and Doob, Leonard W., "Competition and Cooperation," *Soc. Science Res. Council Bulletin*, 1937, p. 191), and the data they reviewed should be integrated by expressing their "goals, persons, rules, performances," etc., in S-notation.

35. The list of aggregated processes might be enlarged and critically revised as a preliminary to testing these definitions upon collections of appropriate data. What combinative functions are best? Will structural formulae (see S. 26, Ch. XI) prove useful?

#### XI. PROBLEMS CONNECTED WITH CHAPTER XI, FORCES

36. Now that forces can be measured wherever the change they produce may be measurable, societal forces and their resisting forces need to be experimentally isolated and studied under specified conditions. (Study Ref. 12.)

37. Similarly, situations involving measured societal control under specified conditions should be collected and inductions of theory derived from them.

#### XII. PROBLEMS CONNECTED WITH CHAPTER XII, APPRAISAL OF S-THEORY

38. Situations whose properties are predicted by their quantic number but which have never been reported in the literature of any of the social

sciences for lack of adequate technics of observation should be "discovered." By adopting an unreported quantic number, such as 8;2;0;2, and devising observational technics under some specified conditions, these unexplored sociological regions should be staked out and developed.

39. For a convenient predictive tool for laymen the whorls of probability ogives as in S. 4, Ch. XII might be calculated and published for correlations ranging from 0 to 1.0 (perhaps at .05-point intervals), expressed in  $\sigma$  (and P.E., and A.D., and percentile) units.

### *XIII. PROBLEMS CONNECTED WITH THE APPENDICES*

40. More complete standardization of S-notation, especially for Brief-S formulae, should be secured. With the aid of controlled experiments, measuring the percentages of agreement of independent analysts, the relative excellence of any ruling to be standardized may be determined.

41. The formulae for the standard errors of sampling should be derived for all indices in this volume for which they are lacking, before their statistical significance can be determined in given data.

42. A further study, after a while, to map out areas for further research studies might be useful in co-ordinating efforts to extend the frontiers of knowledge in this field. (Cf. #13.)

## Appendix IV

### BIBLIOGRAPHY \*

1. Allport, F. H., "The J-Curve Hypothesis of Conforming Behavior," *Journal of Social Psychology*, 1934, 5, pp. 141-183.
2. American Public Health Association, 50 West 50th Street, New York City, *Appraisal Form for City Health Work*, 4th edition, 1934, p. 136.
3. Bernard, L. L., *An Introduction to Social Psychology*, Henry Holt and Co., 1926, p. 652.
4. Beyle, Herman C., *Identification and Analysis of Attribute-Cluster-Blocs*, University of Chicago Press, 1931, p. 249.
5. Cajori, Florian, *A History of Mathematical Notation*, Open Court Publishing Co., Vols. I and II, 1928.
6. Chapin F. Stuart, *Contemporary American Institutions*, Harper and Bros., 1935, p. 423.
7. —, "Social Theory and Social Action," (Presidential Address), *Amer. Sociol. Rev.*, Vol. I, No. 1, Feb., 1936.
8. Christian Century, Editorial, Vol. LIII, No. 17, April 22, 1936, p. 587.
9. Cahen, Alfred, "Probability of Divorce," *Journal of the American Statistical Association*, Proceedings, Vol. XXVII, New Series, No. 177A, 1932, pp. 46 ff.
10. Committee on the Introductory Course, "Report to the American Sociological Society," *Jour. Educ. Sociol.*, Vol. VII, No. 1, Sept., 1933, pp. 80-82.
11. Dewey, John, *How We Think*, D. C. Heath and Co., 1910, p. 224.
12. Dodd, Stuart C., *A Controlled Experiment on Rural Hygiene in Syria*, American University of Beirut, 1934, p. 336.
13. —, "On the Sampling Theory of Intelligence," *Brit. Jour. Psych.*, Vol. XIX, Part 8, Jan., 1929.
14. —, "Progress Inductively Defined," *Internat. Jour. of Ethics*, Vol. XLIV, No. 3, April, 1934.
15. —, "A Social Distance Test for the Near East," *Amer. Jour. Sociol.*, Vol. XLI, No. 2, Sept., 1935.
16. —, "The Standard Error of a Social Force," *Annals of Mathematical Statistics*, Vol. VII, No. 4, Dec., 1936.
17. —, "The Theory of Factors," *Psych. Rev.*, Vol. 35, Nos. 3 and 4, May and June, 1928.

\* Addendum in proof:

Dodd, Stuart C. "A System of Operationally Defined Concepts for Sociology," *American Sociological Review*, Oct. 1939, Vol. IV, No. 5. This is a resumé of S-theory.

18. —, "A Theory for the Measurement of Some Social Forces," *Scientific Monthly*, Vol. XLIII, No. 1, July, 1936.
19. —, "A Tension Theory of Societal Action," *Amer. Sociol. Rev.*, Feb., 1939, No. 1, Vol. III.
20. Dudycha, George, "An Examination of the J-Curve Hypothesis Based on Punctuality Distributions," *Sociometry*, Vol. I, Nos. 1 and 2, July-Oct., 1937.
21. Dunhan, H. W., "The Ecology of the Functional Psychosis in Chicago," *Amer. Sociol. Rev.*, Vol. II, No. 4, Aug., 1937.
22. Dunlap, J. W., and Kurtz, A. K., *Handbook of Statistical Nomographs, Tables, and Formulas*, World Book Co., 1932, p. 163.
23. Eaton, R. M., *Symbolism and Truth*, Harvard University Press, 1925.
24. Engelhart, Max D., "The Technique of Path Coefficients," *Psychometrika*, Vol. I, No. 4, Dec., 1936.
25. Eubank, E. E., *Concepts of Sociology*, D. C. Heath and Co., 1932, p. 570.
26. Freeman, F. N., "Heredity and Environment in the Light of the Study of Twins," *Scientific Monthly*, Vol. XLIV, No. 1, Whole No. 256, Jan., 1937.
27. Gulliksen, H., "The Content Reliability of a Test," *Psychometrika*, Vol. I, No. 3, Sept., 1936.
28. Hankins, Frank H., "Social Discrimination," *Encyclopaedia of the Social Sciences*, Macmillan, Vol. XIV, 1935, pp. 131-134.
29. Hotelling, H., *Analysis of a Complex of Statistical Variables into Principal Components*, Warwick and York, Baltimore, 1933.
30. —, "The Most Predictable Criterion," *Jour. Educ. Psych.*, Vol. XXVI, No. 2, Feb., 1935, pp. 139-142.
31. House, R. N., "Measurement in Sociology," *Amer. Jour. Sociol.*, Vol. XL, No. 1, July, 1934.
32. Hudson, R. G., and Lipka, J., *A Manual of Mathematics*, John Wiley and Sons, 1917, p. 132.
33. Katz, D., and Cantril, H., "Public Opinion Polls," *Sociometry*, Vol. I, No. 1 and 2, July-Oct., 1937, p. 175.
34. Kelley, T. L., *Essential Traits of Mental Life*, Harvard University Press, 1935, p. 145.
35. —, *Statistical Method*, Macmillan, 1923, p. 390.
36. Korzybski, Alfred, *Science and Sanity*, Lancaster, Science Press, 1933, pp. 798.
37. Krause, Ruth H., and Conrad, Hubert S., "An Extension of the Kelley-Wood and Kondo-Elderton Tables of Abscissae of the Unit Normal Curve for Areas between .4500 and .49999 99999. *Jour. Exper. Educ.*, March, 1937.
38. Kuczynski, R. R., "Population—History and Statistics," *Encyclopaedia of the Social Sciences*, Macmillan, 1934.
39. Lasswell, H. D., "Chinese Resistance to Japanese Invasion: The

- Predictive Value of Precrisis Symbols," *Amer. Jour. Sociol.*, Vol. XLIII, No. 5, March, 1938.
40. Leahy, A. M., *The Measurement of Urban Home Environment*, University of Minnesota Press, 1936, p. 70.
  41. Lumley, F. E., *Principles of Sociology*, McGraw-Hill Book Co., 1928, p. 562.
  42. —, *Principles of Sociology*, McGraw-Hill Book Co., 2nd Ed., 1935, p. 461.
  43. Lundberg, George A., *Foundations of Sociology*, Macmillan, 1939.
  44. —, Review of "A Controlled Experiment on Rural Hygiene in Syria," *Social Forces*, Vol. XIX, No. 1, Oct., 1935.
  45. —, *Social Research*, Longmans Green, 1929, pp. 380.
  46. —, "The Thoughtways of Contemporary Sociology," *Amer. Sociol. Rev.*, Vol. I, No. 5, Oct., 1936.
  47. May M., Allport F., and Murphy, G., "Memorandum on Research in Competition and Co-operation," *Social Science Research Council*, 1937, p. 404.
  48. Moreno, J. L., *Who Shall Survive*, Nervous and Mental Disease Publishing Co., 1934, p. 437.
  49. Mukerjee and Sen Gupta, *Introduction to Social Psychology*, D. C. Heath and Co., 1928, pp. 304 and xv.
  50. North, Cecil, *Social Differentiation*, University of North Carolina Press, 1926, p. 343.
  51. Ogburn, W. F., "Stationary and Changing Societies," *Amer. Jour. Sociol.*, Vol. 42, No. 1, July, 1936.
  52. Ogburn, W. F., and Talbott, N. S., "The Measurement of the Factors in the Presidential Election of 1928," *Social Forces*, 8(2), December, 1929.
  53. Ogden, C. K., and Richards, I. A., *The Meaning of Meaning*, Harcourt Brace and Co., 1927.
  54. Park, R. E., and Burgess, E. W., *Introduction to the Science of Sociology*, University of Chicago Press, 1921, p. 1041.
  55. Phelps, Harold A., *Principles and Laws of Sociology*, John Wiley and Sons, 1936, p. 544.
  56. President's Research Committee, *Recent Social Trends in the United States*, McGraw-Hill Book Co., One Volume Edition, 1933, p. 1541.
  57. Reinhardt, J., and Davies G., *Principles and Methods of Sociology*, Prentice Hall, 1932, p. 683.
  58. Reitz, H. L., ed., *Handbook of Mathematical Statistics*, Houghton Mifflin and Co., 1924, p. 221.
  59. Rice, Stuart C., *Statistics in Social Studies*, University of Pennsylvania Press, 1930, p. 222.
  60. Ross, E. A., *Principles of Sociology*, (revised) The Appleton-Century Co., 1930, p. 592.
  61. Semple, Ellen C., *Influences of Geographic Environment*, Henry Holt and Co., 1911, p. 683.

62. Shipman, G., "Science and Social Science," *Social Forces*, Vol. X, pp. 38-48.
63. Simmel, Georg, "The Problem of Sociology," *Annals Amer. Acad. Pol. and Soc. Sci.*, VI, (1895).
64. Smith, J. G., *Elementary Statistics*, Henry Holt and Co., 1934, p. 517.
65. Sociometry, *A Journal of Inter-Personal Relations*, Box J, Beacon, N. Y., Vol. I, 1937.
66. Sorokin, Pitirim, *Contemporary Sociological Theories*, Harper and Bros., 1928, p. 785.
67. —, *Social and Cultural Dynamics*, American Book Co., First 3 vols., 1937, p. 2108.
68. —, *Social Mobility*, Harper and Bros., 1927, p. 559.
69. Spearman, C. E., *The Abilities of Man*, Macmillan, 1927, p. 415.
70. Stephenson, Wm., "The Foundations of Psychometry: Four Factor Systems," *Psychometrika*, Vol. I, No. 3, Sept., 1936.
71. Sturges, H. A., "The Choice of a Class Interval," *Journal of the American Statistical Association*, Vol. XXI, 1926, pp. 65, 66.
72. Thompson, J. R., "The General Expression for Boundary Conditions and the Limits of Correlation," *Proc. Roy. Soc. Edin.*, 1928-29.
73. Thompson, S. P., *Calculus Made Easy*, Macmillan, 1932, p. 301.
74. Thomson, G. H., "Boundary Conditions in the Common-Factor Space, in the Factorial Analyses of Ability," *Psychometrika*, Vol. I, No. 3, Sept., 1936.
75. Thorndike, E. L., and Woodyard, E., "Individual Differences in American Cities: Their Nature and Causation," *Amer. Jour. Sociol.*, Vol. XLIII, No. 2, Sept., 1937, pp. 191-224.
76. Thurstone, L. L., "The Vectors of Mind," *Psych. Rev.*, Vol. XXXI, No. 1, Jan., 1934.
77. —, *The Vectors of Mind*, University of Chicago Press, 1935, p. 266.
78. Von Wiese, L., and Becker, H., *Systematic Sociology*, John Wiley and Sons, 1932, p. 772.
79. Wechsler, David, "The Range of Human Capacities," *Scientific Monthly*, July, 1930.
80. Whitehead, A. N., and Russell, B., *Principia Mathematica*, Cambridge University Press, Vol. I, 2nd Ed., 1925, p. 674.
81. Wissler, Clark, *Man and Culture*, Crowell, New York, 1923, p. 371.
82. Young, K., *Social Psychology*, F. S. Crofts, 1930, p. 680.
83. Yule, G. U., *An Introduction to the Theory of Statistics*, Charles Griffin and Co., London, 1922, Chap. 1, p. 415.

# INDEXES



# AUTHOR INDEX

- Allport, F. H., 121, 250, 297, 923, 925  
 Aristotle, 220  
 Bain, Read, 72  
 Becker, H., 926  
 Bentley, Arthur F., 72  
 Bernard, Jessie, 72, 146, 213, 256  
 Bernard, L. L., 919, 923  
 Beyle, Herman C., 400, 923  
 Billings, Bruce, 72  
 Binnewies, W. G., 624  
 Bogardus, Emory S., 669  
 Bravais, A., 251, 383  
 Bridgeman, P. W., 918  
 Burgess, Ernest W., 121, 862  
 Bushee, Frederick A., 858  
 Cahen, Alfred, 923  
 Cajori, Florian, 923  
 Cantril, H., 924  
 Chapin, F. Stuart, 72, 144, 194, 369, 647,  
     822, 857, 923  
 Child, Irwin, 72  
 Choukas, Michael, 72  
 Conrad, Hubert S., 303, 924  
 Cooley, Charles H., 240  
 Corey, Lewis, 608  
 Cottrell, Leonard S., 862  
 Davies, G., 925  
 Dewey, John, 5, 923  
 Dodd, Stuart C., 130, 131, 326, 364, 367,  
     374, 655, 661, 692  
 Doob, Leonard W., 921  
 Dudycha, George, 250, 297, 924  
 Dunhan, H. W., 924  
 Dunlap, J. W., 924  
 Eaton, R. M., 924  
 Embree, Edwin R., 131  
 Engelhart, Max D., 924  
 Eubank, Earle E., 72, 137, 212, 220, 238,  
     391, 395, 428, 430, 437, 457, 530, 629,  
     707, 735, 834, 894, 896, 916, 924  
 Feldstein, M. J., 647  
 Freeman, F. N., 924  
 Furnas, C. C., 796  
 Garle, H. E., 131  
 Garnett, J. C. M., 326  
 Gauss, 251, 353  
 Gronovsky, A., 649  
 Groves, Ernest, 371  
 Gulliksen, H., 924  
 Hankins, Frank H., 717, 924  
 Himadeh, Said B., 810  
 Hobbes, Thomas, 22  
 Holmes, Samuel J., 372, 665, 702  
 Holzinger, Karl, 72  
 Hotelling, H., 199, 320, 321, 331, 354, 924  
 House, R. N., 924  
 Hudson, R. G., 924  
 Hughes, D. E. R., 361  
 Hull, Clark, 72  
 Jurdak, Angela, 128  
 Kant, Immanuel, 125, 917  
 Katz, D., 924  
 Kelley, T. L., 199, 252, 297, 320, 321, 331,  
     368, 377, 924  
 Kelley-Wood, 924  
 Kondo-Elderton, 924  
 Korzybski, Alfred, 924  
 Krause, Ruth H., 303, 924  
 Kuczynski, R. R., 924  
 Kurtz, A. K., 924  
 Lasswell, H. D., 924  
 Leahy, A. M., 925  
 Leiffer, Murray H., 658  
 Lipka, J., 924  
 Lumley, F. E., 121, 925  
 Lundberg, George A., 3, 12, 21, 29, 72,  
     122, 401, 488, 537, 706, 732, 822, 925  
 McCormick, Thomas C., 373  
 Malik, Charles, 72  
 Marshall, Leon C., 626  
 May, Mark A., 121, 921, 925  
 Mendeleeff, 44; periodic table, 841  
 Menger, Karl, 412  
 Mill, J. S., 313, 384  
 Miller, Herbert A., 240  
 Moore, Elon H., 663

- Moreno, J. L., 389, 393, 401, 416, 670, 925  
Mukerjee, 925  
Murphy, G., 121, 925  
  
Newstetler, W. I., 647  
Nietzsche, 271  
North, Cecil, 925  
  
Ogburn, Wm. F., 370, 371, 383, 724, 925  
Ogden, C. K., 925  
  
Park, R. E., 121, 925  
Pearl, Raymond. See Reed-Pearl  
Pearson, Karl, 299, 313, 725  
Pemberton, Earl, 251, 741, 920  
Phelps, Harold A., 72, 925  
  
Reed-Pearl, 741, 849  
Reinhardt, J., 925  
Reitz, H. L., 925  
Reshevsky, Nicholas, 72  
Rice, Stuart C., 925  
Richards, I. A., 925  
Ross, E. A., 925  
Runner, 401  
Russell, B., 126, 187, 917, 926  
  
Semple, Ellen C., 455, 490, 925  
Sen Gupta, 925  
Shipman, G., 926  
Simmel, George, 391, 926  
Smith, J. G., 926  
Smith, Norman Kemp, 917  
  
Sorokin, Pitirim, 12, 318, 710, 921, 926  
Spearman, C., 199, 320, 322, 926  
Stephenson, Wm., 323, 920, 926  
Sturges, H. A., 926  
  
Talbott, N. S., 383, 925  
Tarde, Gabriel, 579, 921  
Taylor, Griffeth, 375  
Thomson, G. H., 199, 320, 323, 326, 363, 926  
Thompson, J. Ridley, 323, 379, 926  
Thompson, S. P., 926  
Thorndike, Edward, 360, 926  
Thurstone, L. L., 72, 199, 320, 321, 926  
Tibbits, Clark, 370  
  
Von Weise, L., 121, 138, 212, 304, 390, 391, 392, 394, 395, 428, 457, 554, 715, 916, 926  
  
Wallis, 22  
Ward, Lester, 895  
Webster, Daniel, 575  
Wechsler, David, 304, 926  
Weintraub, David, 608  
Wells, H. G., 682  
Weybright, Victor, 366  
Whitehead, A. N., 126, 187, 917, 926  
Wiehl, Dorothy G., 131  
Wissler, Clark, 926  
Woodyard, E., 926  
  
Young, Erle F., 131  
Young, K., 926  
Yule, G. U., 194, 926

# TOPICAL INDEX

- Acceleration, 36, 46; of divorcing, 77, 496; see Celeration, 736
- Accommodating, (Eq. 31, Ch. X) 547, 586
- Accounting equation, of business, (Eq. 24, Ch. X) 541
- Accumulating, 548, (Eq. 32a, Ch. X) 714
- Acting, a qualitative activity, (Eq. 25c, Ch. X) 713
- Actuarial prediction, 864
- Addition, 26, 59; of attributes, 186; qualities, 39
- Adequacy, of sampling, 158
- Adjusting, and maladjusting, (Eq. 93c, Ch. X) 594; science as technic for, 831
- Adjustment, 8; as science, 3, 4; classes of, 396
- Adpopulating, (Eq. 9b, Ch. X) 534; illustrated, 603, 612
- Aesthetic indicators, 151
- Agent, 758
- Aggregated processes, 586, 588; hypotheses, 733
- Aggregation, 26, 30, 157, 212, 899; and units, 913; denoted by, 910; sign, 54, 59
- Aggregative, scripts, 46
- Agreement, a subcriterion, 64; percent of, 66; subcriterion of, 71; testing of, 437
- Algebra, 19
- Alienation, as sine, (Eq. 53, Ch. VI) 342; coefficient of, 383
- All-or-none, indicants, 154; units, 748
- Ameliorating, (Eq. 66b, Ch. X) 569
- Amelioration, (Eq. 29, Ch. V) 259
- Anthropology, 136
- Appetites, see Desiderata, 30
- Areal, class intervals, 50; classes, 47
- Arealizing, (Eq. 67c, Ch. X) 725, 726
- Aristotelian, laws of logic, 24; logic, 298
- Arrays, variance of, 332
- Assimilating, illustrated, 630; limit of, (Eq. 26a, Ch. X) 714
- Associating, (Eq. 6a, Ch. X) 533; illustrated, 601
- Asymmetry, of the interrelational matrix, 103, 405
- Atom, social, 422
- Attensing, definition, (Eq. 35a, Ch. X) 715
- Attenuation, (Eq. 48, Ch. VI) 338
- Attitude tests, 15
- Attracting, 586, (Eq. 91, Ch. X) 589
- Attribute, 39; and length product, (Eq. 5a, Ch. VIII) 452; hypothesis, 39, 185; hypothesis, research upon, 919; implicit, 84, 125, 139; multiplication of, 169; product of, (Eq. 9b, Ch. III) 141, 154, 239, 903; product with population, 207; summed, multiplied, divided, 186; when written explicitly, 238
- Attribute cluster blocs, 400
- Attribute hypothesis, and mathematical rules, 172
- Attribute product, and multiple plurel script, (Eq. 46, Ch. IV) 209; illustrated, 816
- Attribute-space products, 491
- Attribute-time product, (Eq. 8a, Ch. IX) 497; (Eq. 8b, Ch. IX) 524
- Attribute units, 747
- Attributes, aggregate of changing, (Eq. 25e, Ch. X) 714; an aggregate change, (Eq. 25f, Ch. X) 714
- A-unit, 753
- Base letters, 903
- Behaving, (Eq. 29a, Ch. X) 545
- Behavior, function of, (Eq. 32, Ch. V) 264
- Binary, celeration, 736; process, 531; spatial processes, 726
- Binomial expansion, (Eq. 10, Ch. VI) 382
- Bio-cultural indicators, 149
- Biological indicators, 147
- Biological tension, 267
- Bipartite matrix, 406, 913
- Biplaner surfaces, 404
- Bivalent relation, 406
- Blocs, of interrelations, 399
- Borderline, quantic, 175, 176, 183
- Brief S-formulae, 506; product of a quality and a quantity, 527, 529; see Eqs. 21-28, Ch. IX, 526
- Budget, and tension, (Eq. 51, Ch. V) 275; family, 802

- Calendar, world, 508  
 Calibrated, indices, 157; units, 89  
 Capital letter, as scripts, 54; symbolizing indices, 48  
 Capitalism, 259  
 Capitalistic competing, 560  
 Cardinal, 16; indicants, 155; units, 749  
 Cartesian co-ordinates, 319  
 Case script, indicatory 162 ff.; (Eq. 24, Ch. III), 194; rules for, 908; the pre-superscript, 33, 51  
 Cases, 26, 58, 900  
 Categories, defined by quantics, 314; Eubank's, 173; Kant's, 917; non-overlapping, 346; of Eubank and S-theory, 894; of plurals, 212  
 Causal analysis, of a case, 517, 614  
 Causation, a formula for, (Eq. 15, Ch. XI) 822; and correlation, 313; canons of, 384; illustrated, 636, 806-7, 816  
 Cause, and mechanism, 348a, 822; illustrated, 7  
 Causes, vs. forces, 757  
 Celeration, classified by sectors, 740; classified temporally, 737; cycle, (Eq. 3, Ch. XI) 738; definitions, (Eq. 1a, Ch. XI) 736; fluctuant, (Eq. 4, Ch. XI) 739; formulae, 523; standard error of, 823; trending, 737  
 Cell entries, 44  
 Cells, number in quantic table, 45  
 Centroids, reduction to, 321  
 Change, amount of, 25, 46; see Ch. X, 528; definitions, 528; standard error of, (Eq. 19, Ch. XI) 823  
 Chapter sequence, 45  
 Characteristics, 25, 26; and indicators, 153; classes of, 47, 58, 135; definition of, 28  
 Ciphering, 22  
 Class script, 206; examples of, 48; on spatial indices, 452; on the indicator, 144; rules for, 907; the post-subscript, 33, 46  
 Classification, of celeration, 737; of societal processes, 586; standardized, 576  
 Classification script, (Eq. 4a, Ch. II) 126  
 Classifications, 7; cross, 57; experimental verification of, 394; in S-theory, 62; of characteristics, 144; of culture, 191; of indicators by content, 146; of indicators by precision, (Eq. 16, Ch. III) 153; of indicators, refined, 919; of plurals, 211; of population, 55; of sectors, 28; of time, 76, 136; reliability of, 190; reversible, 57; societal change, 530; unsatisfactoriness of, 191  
 Class-interval script, spatial, (Eqs. 3, 4, Ch. VIII) 489, 908; the pre-subscript, 33, 81, 219  
 Class-intervals, 26; best number of, 191, 900; quantitative indicators, 49; script, 153; scripts, 49; spatial, 50, 58  
 Classes, 26, 58, 900  
 Cliques, 399  
 Coercing, 586, 590, (Eq. 91.5b, Ch. X) 591  
 Cohesion, of the group measured, 416  
 Colon, analogue of symbol, 126; as symbol of aggregation, 54; denoting dependence, 56; double, 57, 77, 78; interpretations of, 198; rules for, 910; symbol, 207; unusual subclassification, 228, 900; with meaning verbalized, 55  
 Color, quantified, 39  
 Combination, aggregative, 30, 157; and in groups, 442; and permutation in groups, 412; of attributes, 186; of n-things, (Eq. 10b, Ch. VI) 382; of symbols, 18  
 Commercializing, 587; (Eq. 98, Ch. X), 734  
 Communal, indicators, 152; situations, 302  
 Communality, (Eq. 45, Ch. VI) 338; (Eq. 104, Ch. VI) 352  
 Communication systems, 149  
 Communism, 259  
 Community, challenge to test definitions of, 442; in India, 215; operationally defined, 410  
 Commutative matrix, 97  
 Comparability, of units, 161, 246  
 Competing, and co-operating behavior, 385; equality to monopoly case, (Eq. 55, Ch. X) 721; illustrated, 613, 632, 633, 653, 788; in average deviation terms, (Eq. 47b, Ch. X) 721; (Eq. 57b, Ch. X) 722; index of, (Eq. 47a, Ch. X) 559, 587; measurement of, 95; research upon, 921; special cases, (Eqs. 48-52, Ch. X) 718  
 Complex processes, 531, 577  
 Complexity, of data handled by matrices, 156  
 Components, analysis of general, 162, 320, 326; compulsory, 322; interrelational, 409, 759; purpose of, 353; reduction, 320; technics of analysis, 733; units of, (Eq. 102, Ch. VI) 352

- Composition, of forces, 754  
 Compound, change, definition of, 530  
 Comprehensiveness, of S-concepts, 834  
 Concept, 6; added to S-theory, 918; disagreement upon, 24; eventually defined, 919; number of basic, 895; number of in S-theory, 18; of S-theory, 60; S-formulae for, 69, 833  
 Conflicting, (Eq. 10, Ch. X) 535, 586; illustrated, 665  
 Consistency, internal, 64  
 Contact, combinations, 412; index of, (Eq. 13, Ch. VII) 440; matrix, (Eq. 12b, Ch. VII) 439; societal, (Eq. 12a, Ch. VII) 405  
 Contacting, illustrated, 669, 670  
 Contingency, 312, 336; and interrelation, 440; coefficient of, (Eq. 51, Ch. VI) 341; illustrated, 642  
 Continuum, interrelational, 397  
 Control, 46; as function of science, 3; evaluatively neutral, 762; illustrated, 811; societal, (Eq. 16a, Ch. XI) 760; to be studied, 921  
 Co-operating, 548, 586; illustrated, 633, (Eq. 33a, Ch. X) 714  
 Co-ordinates, oblique, 384  
 Correlates, 142  
 Correlating, 26, 59  
 Correlation, 45; analysis of, 319; and causation, 313, (Eq. 15a, Ch. XI) 822, 900; and component analysis, (Eq. 100, Ch. VI) 352; and cosine, (Eq. 52, Ch. VI) 342; as a product of probabilities, 311; as the cosine, 316; biserial, (Eq. 36, Ch. VI) 333; coefficient, (Eq. 3, Ch. VI) 308, 329; curvilinear, (Eq. 35, Ch. VI) 332; dependent on range, (Eq. 92, Ch. VI) 350; double multiple, 354; in geometric terms, 316; in terms of probability, 310; interpreted by diagrams, 328; multiple, (Eq. 113, Ch. VI) 354; notation for, 127, 329; notation of, 356; of a sum and a part, (Eq. 13b, Ch. VI) 386; of ranks, (Eq. 87, Ch. VI) 348; part, 356; partial, (Eq. 117, Ch. VI) 355; Pearson product-moment, 40; ratio of, (Eq. 4, Ch. VI) 380; scattergram, 57, (Eq. 1, Ch. VI) 308, 381; spatial, (Eq. 41, Ch. VI) 335; surface, (Eq. 39, Ch. VI) 334; triple, 376; variant forms, 380-1  
 Correlation coefficient, 911, 913  
 Correlation mechanisms, 352  
 Correlation scattergram, 913  
 Correlation sign, 57  
 Correspondence, criterion of, 70; external, 64; symbol, 77  
 Cosine, and interrelationship, 391; as correlation, (Eq. 52, Ch. VI) 118, 342  
 Cost of production, theories of, 920  
 Covariances, defining-processes, 563  
 Cranial capacity, formulae, (Eq. 18a, Ch. IX) 504, 506  
 Criteria, for frequency curves, (Eq. 24, Ch. V) 299  
 Cross-classification, 26; formula for, (Eq. 5, Ch. VII) 436; interrelations, 390; rules for, 911; sign, 57, 59; table, (Eq. 1, Ch. VI) 308  
 Cross-scripting, 238, 506  
 Crowds, (Eq. 10b, Ch. VIII) 458; formulae for, (Eq. 11, Ch. VIII) 491; photographic example of, 488  
 Cubic, class-intervals, 50  
 Culture, base, coupled, traits, 156; complexes, indicators of, 526; example of, 732; form of descriptive formula for, 144; indicators, 147; rate of diffusion of, 109  
 Culture areas, 462; example of, 464  
 Culture changes, differential velocities of, (Eq. 62b, Ch. X) 724  
 Culture traits, hypothesis of, (Eq. 86a, Ch. X) 583; longevity of complexes, (Eq. 15, Ch. IX) 501  
 Cumulating, 549, (Eq. 32, Ch. X) 714  
 Customs, 502  
 Cyclic celeration, (Eq. 3, Ch. XI) 738  
 Data, need for better, 215  
 Date script, 52, 93, 498, (Eq. 11a, Ch. IX) 525  
 Date, symbol for one, 58  
 Dateless situations, 36  
 Dates, aggregation of, (Eq. 1c, Ch. IX) 522  
 Deceleration, see Celeration, 736  
 Decumulating, 548, (Eq. 32b, Ch. X) 714  
 Deduction, 9  
 Definitions, of sociology, 12; operational, 156; see Operational, 10  
 Degree, of equations, 125, (Eq. 36, Ch. III) 199; of matrices, 192; of matrix, (Eq. 5, Ch. IV) 239; of the matrix, 430, 910  
 Densating, 574, 726  
 Densing, (Eq. 68, Ch. X) 574, 587, 728  
 Densities, see Ch. VIII, 449; classification of, 453  
 Densitive situations, 302

- Density, 46; and culture, 920; and productive-culture, (Eq. 8a, Ch. VIII) 455; defined, (Eq. 7a, Ch. VIII) 453; inverted, (Eq. 7b, Ch. VIII) 454; of plurels, (Eq. 9a, Ch. VIII) 490; of plurels vs. plurels of, 457
- Depopulating, (Eq. 9c, Ch. X) 534; illustrated, 606, 612
- Descriptive formula, as stating findings, 281; equations, 18, 59; reliability of, 66
- Descripts, 46, 157; rules for, see Appendices; temporal, 497
- Desiderata, 29, 252 ff., (Eq. 25c, Ch. V) 253, 263; special cases of, 271; systems of, 273; unlimited, (Eq. 42, Ch. V) 270
- Desire, 263; indices of, (Eq. 35, Ch. V) 265, 305
- Detensing, definition of, (Eq. 35b, Ch. X) 715
- Determinant, of intercorrelations, 352
- Detracting, (Eq. 91, Ch. X) 589
- Diagrams, interpreting correlation, 328
- Differentiation, 145; of plurels, (Eq. 9, Ch. IV) 218
- Differing, 544, (Eq. 25b, Ch. X) 713; illustrated, 621, 628
- Dimensional analysis, 319, 918
- Dimensions, example of, 197; examples of transformation, 320; exponential, 736; geometrically defined, 165; geometry of, 67; illustrated, 638; in a recorded societal situation, 840; in algebra vs. geometry, (Eq. 36, Ch. III) 198; number and kind, 48; number in the economic process, 712; number of, 68, 320; number of, (Eq. 35a, Ch. III) 165; number of in a situation, 235, 430; number of necessary, 199; number of populational, (Eq. 19, Ch. IV) 224; number of spatial, (Eq. 6, Ch. VIII) 453; oblique, 199; of societal change, 586; of the moments, 243; of time, 495; orthogonal, 319; parsimony of, 172; physical and statistical, 199; physical vs. statistical in the space sector, 489; societal, 15, 16; temporal geometric interpretation of, 525; vectorial, illustrated, 638
- Diphthong, 84; in index, 210
- Direction cosines, (Eq. 75, Ch. VI) 345-6
- Discriminating, 717
- Dispersing, (Eq. 45a, Ch. X) 557
- Dispersion, graphing of, 289; maximal, 95, 247 (Eq. 21b, Ch. V), 249; measure of, (Eq. 9, Ch. V) 243; of human characteristics, 303; relation to comparability of units, 246; relation to probability, 245
- Dissimilarity, 145, (Eq. 21, Ch. V) 249
- Dissimilarizing, illustrated, 613, 626, 633, 645
- Dissociating, (Eq. 6b, Ch. X) 533; illustrated, 612
- Distributing, 726
- Distribution, 45; a list of persons, 913; frequency, 56; line of equal, 284; list vs. frequency, 242; of indicants, 242; of persons, 913; of persons by ages, 913; of persons frequency, (Eq. 1a, Ch. V) 242; of plurels by size, (Eq. 4, Ch. V) 294, 913; terms for sections of, 254
- Division, of attributes, (Eq. 7, Ch. III) 187; qualitative, 29, 39, 59
- Divorce, probability of, (Eq. 15, Ch. IV) 221
- Domestic, indicators, 151, 256; situations, 300
- Dot, heavy, 40; meaning of heavy, 57, 900, 911
- Double colon, 900
- Drives, see Desiderata, 30
- Durating, 36, 575, 586, (Eq. 71a, Ch. X) 729
- Duration, 34, 35
- Durational, recodispersing, (Eq. 76, Ch. X) 731; recovering, (Eq. 77, Ch. X) 731; redispersing, (Eq. 72, Ch. X) 731; reranking, 731; revarying, (Eq. 73, Ch. X) 731
- Durational populating, 586, 730
- Durational processes, 574
- Durations, see Ch. IX, 495
- Dyad, 399; relation of, 390
- Dynamic, indicants, 101; situations, 36
- "E," as elements, 183; as exponent, 186
- $\epsilon$  (epsilon), elements, 162
- Earning, process of, 35
- Ecological, matrix, 47
- Economic, 136; indicators, 151, 256; production factors, 185; situations, 300
- Economic process, (Eq. 23a, Ch. X) 540: formula of, (Eq. 23b, Ch. X) 712
- Economic tension, 269
- Economic value, (Eq. 41, Ch. V) 269
- Educational, indicators, 151, 257; situations, 301; tension, 267
- Effective force, 743
- Effectiveness, of coercing, 591
- Elements, as sociological atoms, 156; number of common, (Eq. 19b, Ch. VI) 387; percent of common, (Eq. 42, Ch.

- VI) 336; percentage of common, (Eq. 13a, Ch. VI) 324; properties of, 327
- Ellipses, of correlation, (Eq. 28, Ch. VI) 331
- Energy, societal, (Eq. 34, Ch. XI) 827
- Engel's law, 659
- Entropy, a possible parallel, 299
- Environments, 25
- Epsilon elements, hypothesis, 920; hypothesis of, 324; interrelational, 409; properties of, 920
- Equalitarian, interrelation surface, (Eq. 16, Ch. VII) 442
- Equalization, (Eq. 28, Ch. V) 258
- Equalizing, (Eq. 45b, Ch. X) 557
- Equation, calculative, 59; calculative example, 233; descriptive, 59; references to, 124; uses of, 222
- Equilibrium equation, 265
- Equipment, office, store, etc., 149
- Equiproportional, (Eq. 20, Ch. VI) 326
- Equivalence, 901
- "Eq.," 185
- Error, analyzed, 193; constant, (Eq. 21, Ch. III) 159; observand, 159; observee, 159; observer, 159; synthesis of, 338; variable, (Eq. 22, Ch. III) 159; of measurement, (Eq. 46, Ch. VI); variance of, 338
- Esthetic indicators, 257
- Ethico-religious indicators, 151, 257; situations, 302
- Ethics, 137
- Euclidean geometry, and other, 918
- Evaluating, illustrated, 652
- Event, a qualitative action, 914
- Evolution, formula for, 733
- Evolution-to-normal hypothesis, 252
- Evolving, 587, (Eq. 97a, Ch. X) 597
- Exchange matrix hypothesis, 542
- Experiment, crucial, 9; testing S-theory, 65
- Experimental testing, of a classification, 394; of objectivity of concepts, 833
- Exploiting, 587, (Eq. 96, Ch. X) 596
- Exponent, 26; and tension theory, 121; indicatory, 38, 58, (Eq. 2, Ch. III) 139, 157; meaning of, 124, 901; of physical space, 451; of unity understood, 34; populational, 37, 203; positive and negative, 76, 107; rules for, 904; spatial, 33; temporal, 34, 496, 899; the post-superscript, 33
- Exposition, two levels of, 23
- External correspondence, criterion of, 70
- Facilitating force, 754
- Fact, definition of, 6
- Factor, 320, 758
- Families, types of expressed in S-notation, (Eq. 14, Ch. IV) 221
- Feelings, see *Desiderata*, 30
- Financial statement, S-formula for, 713
- Fluctuant celeration, (Eq. 4, Ch. XI) 738
- Folkways, 502
- Force, 36, 46; aggregation of, illustrated, 790; as a vector, 745; definitional assumptions, 744; economic attensing, 802; example of economic, 742; example of educational, 742; illustrated, 777, 790, 799; measurability of, 821; net societal, 77; of public opinion, 742; per capita, (Eq. 12, Ch. XI) 751; per capita illustrated, 796, 797; physical, (Eq. 8a, Ch. XI) 821; problem of units, illustrated, 800; social, 99; societal defined, (Eq. 7, Ch. XI) 742; standard error of, 759, (Eq. 16, Ch. XI) 823; vs. a cause, 757
- Forces, see Ch. XI, 736; compounding of, illustrated, 784; of business, (Eq. 9, Ch. XI) 746; summation of, illustrated, 786; systems of, 753; to be experimentally isolated, 921; units for, 746
- Formula, function of, 502, 545
- Formulae, function of, 505
- Foundations of Sociology*, 3, 12, 16
- Frequencies, 586
- Frequency distribution, 7
- Fruitfulness, of S-theory, 63
- Function, 126
- F-unit*, 753; illustrated, 788
- G and s factors, 322
- Gaussian curve, 383
- Generalization, 6, 7; reason for, 8
- Geometric system, 839
- Geometry, of indicators, 162 ff.; of populational sector, 223; of S-theory, 67
- Gestalt theory, 756
- Goodness-of-fit, of interrelational surfaces, 412
- Grading, (Eq. 34, Ch. X), 549; connotation, 714
- Group-cultural indicators, 150
- Groups, borderline situation, 429, 441; contact combinations, 412; defined, 389; differentiated from plurels, 919; interrelational combinations, 414; political classification of, (Eq. 13, Ch. IV) 220; primary vs. secondary, 240; quantic formula, (Eq. 1, Ch. IV) 205,

- 212; sociological, 38; structures, 442, (Eq. 10c, Ch. VIII) 458  
 Growth, denotation of, 528
- Hatred, quantified, 39  
 Hectare, 490  
 Hero, 400  
 Heterosectoral, indices, 189  
 Hierarchical differing, (Eq. 25d, Ch. X) 713  
 Hierarchy, notation for, 913; of attribute population products, (Eq. 4c, Ch. IV) 210; of characteristics, (Eq. 13, Ch. III) 144; of periods, 55; of persons and plurals, (Eq. 7, Ch. IV) 216; of purels, 55, (Eq. 3a, Ch. IV) 206; of subclasses, 86; visualized, (Eq. 36, Ch. IV) 207  
 Homosectoral, indices, 189; scripts, 49  
 Homosectoral index, 32; rule for, 904; symbol for, 58; (Eq. 28, Ch. II) 125; 899  
 Household equipment, 149  
 Human category, (Eq. 10a, Ch. VIII) 458  
 Hypotheses, for aggregate processes, 733; in S-system, 842; of imitation, (Eq. 80, Ch. X) 579; to be tested, 916  
 Hypothesis, indicant, 141; e-hypothesis, 183; J-curve, 250; of a natural range, 259, 920; of epsilon elements, 324, 920; of evolving toward the normal, 252; of formulae, 546; of multiplex-elements, 251; of normal pyramids, 107; of reliability in science, 761; of tension, 554; of the exchange matrix, 542; underlying this volume, 3; working, 121
- I, see indicator, 31  
 Identification, 26; script, (Eq. 49, Ch. II) 127; sign, 58  
 Identity, 901  
 I-force, (Eq. 11, Ch. XI) 751; illustrated, 772, 773, 776, 778, 780, 782, 813, 814  
 Imitation, 579; hypotheses to be tested, 921  
 Implicit, index, 188  
 Importance, of coercing, 591  
 Income, distributions of, 645  
 Indefiniteness, of indices, 490, 188; symbol of, 84  
 Index, formula for, (Eq. 10, Ch. III) 188, 899; implicit, indefinite, and nul, 188; rule for, 904; statistical, 7; symbol for, 32; symbol for homosectoral, 58, 142, (Eq. 10a, Ch. III) 157  
 Indexing, (Eq. 3b, Ch. X) 530, 543, 586  
 Index number, 143, 161, (Eq. 11a, Ch. III) 189  
 Indicants, 141; degrees of precision of, 154  
 Indicating, 544, 586; illustrated, 649, (Eq. 28, Ch. X) 714  
 Indicator, correlated, 142; defined, 145, 899; definition of, 28; exponent on, 38, 45, 135; rule for, 903; change, (Eq. 27, Ch. X) 714; symbol for, 31  
 Indicatory, case script, 52, 162; class-interval, 153; class script, 47, 144; densities, 462; durations, 501; exponent, (Eq. 2, Ch. III) 139, (Eq. 12, Ch. III) 143; exponent, rules for, 905; formula for, (Eq. 1, Ch. III) 136; interrelations, 395; processes, 543.  
 Indices, 26; homosectoral, and heterosectoral, 189; summative, 30, 157  
 Individualizing, (Eq. 93b, Ch. X) 594  
 Inequality, judgments of, 91, 901  
 Inheritance, social, 584  
 Insight, in research, 22; into relationships, 157  
 Institutional, groups, 151  
 Institutions, type parts, 97, 144  
 Insurance, 500, 921  
 Integrating, 586, 735  
 Intensity, of desire, (Eq. 31, Ch. V) 263  
 Intentional, defined, 714  
 Interacting, (Eq. 12a, Ch. X) 537  
 Interaction, defined, 389; matrix, illustrated, 672; relationships, 396  
 Intercorrelations, as direction cosines, 317; matrix of, (Eq. 43b, Ch. X) 717; symbolized by, 912  
 Interests, see desiderata, 30, 151  
 Interference, in imitation, (Eq. 82, Ch. X) 581  
 Interrelational combinations, 414  
 Interrelation matrix, 88, 103, 205, (Eq. 1, Ch. VII) 389; as completer observation, 392; as operational definition, 393; isolating one variable, 392  
 Interrelations, 46; a geometric interpretation, 436; and contingency table, 440; as a problem, 435; classified by content, 394; classified by surface of the matrix, 398; classified by the form of internal structure, 412; complete formula for, (Eq. 4, Ch. VII) 435; definition, 389; factorable, 409; illustrated, 668; quantic formula for, 390; sociological importance, 391; static vs. dynamic, 708; summarized in one index, 390; symmetric, 405

- Intolerating, (Eq. 92c, Ch. X) 593  
 Inverse, and observe densities, 454  
 Isolation, contacting distinguished, 708;  
   forms and consequences of, 404; in  
   plurels, (Eq. 10b, Ch. VII) 438; so-  
   cietal, (Eq. 10a, Ch. VII) 403  
 IQ's, in S-notation findings, 222  
 I-units, 142
- J-curve, Allport's hypothesis, 250; crite-  
   rion of, (Eq. 23, Ch. V) 298
- Kinship groups, 411  
 Kurtosis, 125; measure of, (Eq. 11,  
   Ch. V) 243, 249, (Eq. 22a, Ch. V) 297
- L, see Length  
 Lag, and lead, 565; by types of culture  
   traits, (Eq. 86b, Ch. X) 584; interval,  
   (Eq. 90, Ch. VI) 349  
 Law, in social sciences, 302  
 Laws, scientific, 7  
 Laws of motion, 822  
 Lead and lag, (Eq. 62, Ch. X) 565  
 Leadership, 400; equations for, 920;  
   formula, defining, (Eq. 8, Ch. VII)  
   401, 420, 438  
 "Legend," abbreviation of, 92, 94; func-  
   tion of, 80, 598; in S-formulae, 73; re-  
   duced, 169; simplified, 281  
 Length, symbol for, 32, 899  
 Letters, to use as scripts, 47  
 Liberating, 592  
 Limit, of indices, 158; lower of a matrix,  
   31, 196, 899; symbolized by, 52  
 Limitation, of S-theory, 842; of S-theory  
   to the quantitative, 124  
 Limits, of sociation, (Eq. 8, Ch. IV) 217  
 Linear class-intervals, 50  
 Linear durating, illustrated, 704  
 Linearizing, 726  
 Lineate densities, 463  
 Lineating, illustrated, 673, 726  
 Linguistic, indicators, 152; situations,  
   302  
 Lining, (Eq. 67b, Ch. X) 725, 726  
 List-distribution, of persons, (Eq. 1b,  
   Ch. V) 293; of plurels, (Eq. 1c, Ch. V)  
   293  
 Localizing, 726  
 Logic, symbolic and S-notation, 187
- Machines, 149  
 Majorates, (Eq. 26d, Ch. V) 254  
 Maloperating, 549, (Eq. 33b, Ch. X) 714  
 Malthusian theory, 115
- Map of areas with lines, 914  
 Mapped regions, rule for, 904  
 Marginal utility, theories of, 920  
 Mathematics for sociologists, 916  
 Matrix, 19; and submatrices, 509; bi-  
   partite, (Eq. 14a, Ch. VII) 406, 432;  
   degree and order of, S1, 193; degree  
   and order illustrated, 638; degree of,  
   107, 172, 910; directional flow within,  
   713; example of, 76; in S-notation, (Eq.  
   19, Ch. III) 157; monogamous, 407;  
   notation in, 126; of indicants, 156;  
   of rank one, (Eq. 20, Ch. VI) 326;  
   order and degree of, 912; order of, 910;  
   polygamous, 407; sectioned, 48; sym-  
   bol for, 32, 46  
 Matrix algebra, and interrelations, 920;  
   exposition of, 192, 438; uses of, 439  
 Maximates, (Eq. 26e, Ch. V) 254  
 Maxinormals, (Eq. 26e, Ch. V) 254  
 Mean, arithmetic, (Eq. 8, Ch. V) 243,  
   244; general formula, (Eq. 12, Ch. V)  
   295, 899  
 Means, 586  
 Measurement, 24; improvement in, 160,  
   193; research on precision of, 919  
 Mechanism, 822  
 Mediates, (Eq. 26c, Ch. V) 254  
 Medical, and vital situations, 300; in-  
   dicators, 151, 256  
 Membership, overlapping in groups, 337  
 Methodology, 18, 216; concepts, in a  
   systematic sociology, 316; hypotheses,  
   and notation of S-theory, 463; problem  
   of observation, 404  
 Migrating, formula of, (Eq. 22a, Ch. X)  
   712; illustrated, 664  
 Minimal, 401  
 Minimality, examples of, 300; types of,  
   256  
 Minimates, (Eq. 26a, Ch. V) 254  
 Mininormals, (Eq. 26a, Ch. V) 254  
 Minorates, (Eq. 26b, Ch. V) 254  
 Mm-unit, 753  
 Mobility, 398; derivation of formulae,  
   708; equality to monopoly case, (Eq.  
   56, Ch. X) 721, (Eq. 57a, Ch. X) 722;  
   illustrated, 612, 663; in average devia-  
   tion units, (Eq. 16a, Ch. X) 710; index  
   of gross, (Eq. 14b, Ch. X) 539, 587;  
   index of net, (Eq. 14a, Ch. X) 539;  
   matrix, (Eq. 13, Ch. X) 538; static  
   aspect, 437; to be tested, 921  
 Mobility formulae, example of use, 710  
 Moment, of any order, (Eq. 6, Ch. V) 294  
 Moments, and correlation, (Eqs. 93-99,

- Ch. VI) 351; matrix of second, (Eq. 43a, Ch. X) 717; of distributions, 242; processes defined by, 570
- Momentum, of a societal change, (Eq. 78a, Ch. X) 578; sample of literacy, (Eq. 79, Ch. X) 579, 587, 732; standard error, (Eq. 88, Ch. X) 732, (Eq. 16, Ch. XI) 823
- Money, 161
- Monoplanar interrelation surface, (Eq. 15, Ch. VII) 442, 408, 417
- Monopolistic trend, illustrated, 788
- Monopoly, dispersion in, (Eq. 19a, Ch. V) 297; index of, (Eq. 20, Ch. V) 248; its formula, 95
- Mores, see Desiderata, 30
- Mortality, process of, 35
- Motivations, see Desiderata, 30
- Multiple, descriptor, 157; plurel script, 206
- Multiple correlation coefficient, 912
- Multiple-membership, technics for dealing with, 710
- Multiplex elements, hypothesis, 251
- Multiplication, of attributes, (Eq. 6, Ch. III) 187; qualitative, 39
- Multiplying, 26, 59
- N-dimensional space, 840
- Negative exponent, (Eq. 1, Ch. IV) 238
- Newton's laws of motion, 743
- Non-determination, coefficient, (Eq. 14, Ch. XI) 756
- Normal correlation surface, (Eq. 39, Ch. VI) 334
- Normal curve, 109, 251, (Eq. 15, Ch. V) 295; illustrated, 797
- Normal ogives, fitted, 920
- Normals, (Eq. 26c, Ch. V) 254
- Norms, 244
- Notation, as methodological hypotheses, 463; of matrix algebra, 19; of S-theory, 329
- Nul, population, (Eq. 1b, Ch. IV) 204; index, 126, 188, (Eq. 1a, Ch. IX) 522
- Nullary, 41; indicatory processes, 544; indices, 185; populations, (Eq. 4a, Ch. X) 204, 531; processes, 570, 586; temporal processes, 575
- Objectivity, of S-concepts, 833
- Observation, 5
- Operating, 549, (Eq. 33, Ch. X) 714
- Operational, 6, 12; definitions, 6, 10, 156
- Operational definitions, limitations of, 725; not achieved, 894; of force, 743; of interrelations, 393; of societal control, 761; of sociology, 12, 181; trial of, 921
- Operational meaning, of the exponent, 496
- Operational rules, for combining attributes, 919
- Operationally defined data, 919
- Operations, sequence of, 723
- Operators, 54; rules for, 909
- Opposing, 587, (Eq. 95, Ch. X) 596
- Order, of a matrix, 193, 430, 910
- Ordinal, 16; indicants, 154; numbers, 104; scales, 15; units, 748
- Ordinating, (Eq. 61b, Ch. X) 724
- Ornaments, 147
- Overlining, denotation of, 132, 900; denoting a vector, 912
- P, see Persons, 32
- P<sup>2</sup>, explained, 390
- Parallelogram, of vectors, (Eq. 62, Ch. VI) 343-4
- Paraphrasing, of sociological concepts, 69
- Parentheses, denotation of, 77
- Parenthesis, meaning of, 912
- Parsimony, furthered by symbols, 51, 60; of dimensions, 172, 320; of S-concepts, 834; of symbols, 22
- Partial correlation coefficient, 912
- Participant, predictions, 844
- Particularity, sign for, 58
- Party, 203
- Path coefficients, 347
- Pathology, social, 255
- Patterns, analyses of societal, 733; tool for measuring, 58
- Peaks, example of, 422; on the interrelational surface, 399
- Pearl, 849
- People, 25, 203
- Peopling, (Eq. 3a, Ch. X) 530, 532
- Per capita index, 37
- Percentages, 161; derivation of, (Eq. 17, Ch. III) 192, 748
- Percentiles, notation for, 913; S-notation for, 296
- Percept, 6
- Period scripts, 509; variants of, (Eq. 9a, Ch. IX) 524
- Periodizing, 575, 586; celeration of, (Eq. 6a, Ch. XI) 821; the formulae for, (Eq. 70a, Ch. X) 729; illustrated, 666
- Periods, of the past, (Eq. 16, Ch. IX) 503

- Permutations, of  $p$  parties, (Eq. 20, Ch. VII) 444
- Persecuting, 587, (Eq. 99, Ch. X) 734
- Person, group matrices, 411
- Person script, 53, 93, 219, (Eq. 12, Ch. IV) 240
- Personal-cultural, indicators, 150
- Personality, 144
- Persons, symbol for, 32; symbol for one, 58, 203
- Persuading, 592
- P-forces, 747; illustrated, 764, 765, 767, 818, (Eq. 10, Ch. XI) 822
- Philosophies of tension, (Eq. 43, Ch. V) 270
- Physical, indicators, 146; dimensionality of space, 489
- Physico-cultural indicators, 147
- Plan, Five-Year, 114
- Planes, of interrelational surface, 404
- Plateaus, of the interrelational surface, 404
- Plurel, 38; and the quantic classification, 205; script, 206
- Plurelizing, 586
- Plurels, 45 ff., 203; and the quantic classification, 205; by size, 237; classification of, 211; differentiated from groups, 919; of density, (Eq. 9b, Ch. VIII) 490
- Plurels-by-size, distribution of, 617
- Plurels of density, 457
- Point scripts, spatial, (Eqs. 1-3, Ch. VIII) 488
- Political, and legal situations, 301; indicators, 151, 256; science, 136; tension, 266
- Polyplanar surface, 411
- Populating, (Eq. 9a, Ch. X) 533; components of, (Eq. 11, Ch. X) 536, 586
- Population, 26; attribute product, (Eq. 4a, Ch. IV) 208; definition of, 27, 58; implicit and nul, 239, 899; indices, 203; nullary, primary, and secondary, 204; rule for, 904; sector, 201 ff.
- Population pyramid, (Eq. 12, Ch. IX) 498; inferences from, 499
- Population score matrix, (Eq. 119, Ch. VI) 357
- Population sector, class-interval, 51, (Eq. 7, Ch. IV) 216, 306
- Populational, case script, 52; class-intervals, 51; densities, 454; during, 587, 730; durations, 498; exponent, rules for, 906; exponent, 203 ff., 219; interrelations, 398; processes, 532
- Populationless situation, 37
- Post-subscript, 46, 206
- Postulates, 3; Lundberg, 16; underlying change, 706
- Power, societal, (Eq. 35, Ch. XI) 827
- Practice, improvement in  $S$ -analysis, 66
- Precision, of  $S$ -concepts, 833
- Precision in measurement, of  $S$ -theory, 61; theory of, 14
- Predicting, as function of science, 3; by calibration, 851; by classifying conditions, 851; by intensity of desire, (Eq. 4, Ch. XII) 853; by multiple correlation, 852; by projecting a trend, 875; by stabilizing the indices, 852; illustrated, 856; in  $S$ -theory, 844; principles for, 850; varying with immediacy, 853
- Prediction, accuracy increasing with subclassification, (Eq. 3, Ch. XII) 895; accuracy of, (Eq. 2, Ch. XII) 850; by distribution curves, 884; dynamic, 849; formula for, 894; from a type, 865; from correlation, 330, 354, 847; from joint probability, 866; from sum of probabilities, 846; from typical temporal pattern, 868; nul cells, 156; participant, 844; probability vs. reliability, 895; static, 845; technic for, 862
- Prestige, ratio of, (Eq. 81, Ch. X) 580
- Presuperdescript, 162
- Primary, 41; indicatory processes, 544; indices, 185; populations, 204, (Eq. 4b, Ch. X) 531; processes, 570, 586; temporal processes, 575; vs. secondary, 411
- Prime, meaning of symbol, 58, 900
- Primed script, 206
- Principia Mathematica*, 126
- Principle axes, 321, (Eq. 29, Ch. VI) 331
- Pre-subscript, 49; indicatory, 51, 153
- Probability, alternative, (Eq. 9, Ch. VI) 382; and causation, 313; joint, 311; laws of, 111; normal, 295; ogives to be calculated, 922; of being normal, (Eq. 14, Ch. V) 245; of coercing, 592; of generalized elements, 327; simple, 310; temporal, (Eq. 7c, Ch. VI) 381; whorl, 341
- Problem, in scientific method, 5; social, 256
- Problem-solving, Dewey's steps, 5
- Process, deductive variation of, 707; denotations of, 528; Von Weise's formula for, (Eq. 38a, Ch. X) 554; Von Weise's

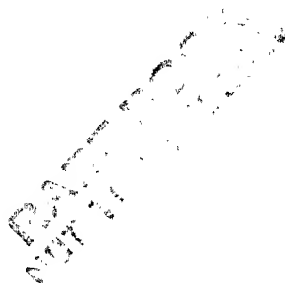
- formula for discussed, (Eq. 38b, Ch. X) 715
- Processes, 46; by form not content, 571; complex, 577; definition of aggregated, 588; Eubank's list of characteristics, 706; formula for, 304; nomenclature, 716; research on aggregated, 921; summary of the single, 585; terminology 530
- Product moment, 307
- Progress, (Eq. 27, Ch. V) 258; denotation of, 528
- Progressing, (Eq. 34a, Ch. X) 549; illustrated, 613, 633
- Prolonging, (Eq. 71b, Ch. X) 729
- Proselytizing, (Eq. 92a, Ch. X) 593
- Puncting, (Eq. 67a, Ch. X) 725, 726
- Pure, indicant, (Eq. 9a, Ch. III) 141; population, 207 ff.
- Purpose, of human life, 831; of this volume, 3
- Pyramid, population, (Eq. 12, Ch. IX) 499
- Q-technic, in correlation analysis, 358; to be explored, 920
- Qualitative, 15; combination, 779; desideratum, (Eq. 48, Ch. V) 273; quantitative product, 527
- Qualitative change, illustrated, 622
- Qualitative product and qualitative ratio, (Eq. 12a, Ch. VIII) 491
- Quality, 39, 45; and nullary processes, 569; as directions, 67
- Quantic, formula, 40; hypothesis, 41; number, 40; "solid," 69, 117; system, 839; Table, 69
- Quantic classification, 44; reliability of, 65, 736
- Quantic formula, categories defined, (Eq. 11, Ch. VI) 315; negative digits in, 125
- Quantic number, 44; illustrative S-situation, 73; in an aggregated situation, 794; of unrecorded situations, 45; rules for, 906
- Quantic solid, sequences of the topics, 437, 841
- Quantitative, concept of, 14, 15; indicators, 139, 140; multiplication, (Eq. 6, Ch. IV) 239; under development of, 17
- Quantity, 39, 45; and primary processes, 659; as directions, 67
- Quaternary, process, 531
- Question mark, 900
- Quinary, process, 531
- Racial, indicators, 152, 256; situations, 301
- Range, hypothesis of a natural, 259, 302; hypothesis upon, 920; in correlation, (Eq. 92, Ch. VI) 350; normally distributed characteristic, (Eq. 30, Ch. V) 261
- Ranks, 161; of the matrix, 199, 748; symbolized by, (Eq. 18a, Ch. III) 192
- Ratio correlation, 332
- Realistic societal analysis, 576
- Re-arealing, 726
- Recodispersing, 563, 568; illustrated, 662, (Eq. 60a, Ch. X) 723, 809
- Reco-ordinating, 564, 568; illustrated, 642, 643, 662, (Eq. 61a, Ch. X) 723-4
- Recorrelating, (Eq. 42, Ch. X) 555, 568; and redeviating, 717; limits of, 725
- Recovarying, (Eq. 89, Ch. VI) 349, (Eq. 65, Ch. X) 567, 568, 657; celeration of, (Eq. 6c, Ch. XI) 821; illustrated, 635, 637, 805, 807
- Recreational, indicators, 151, 257; situations, 301
- Redeviating, (Eq. 41, Ch. X) 555, 568; and recorrelating, 717; limits of, 725
- Redispersing, (Eq. 45, Ch. X) 557, 568; and recodispersing, (Eq. 59, Ch. X) 723; and reordinating, (Eq. 45c, Ch. X) 718; example of, 893; illustrated, 656
- Redurating, 576, 587
- Reference points, 155
- Refraction, in imitation, (Eq. 83, Ch. X) 581
- Regions, 47
- Regrading, (Eq. 57b, Ch. X) 722
- Regressing, (Eq. 34b, Ch. X) 549; illustrated, 613
- Regression, 329
- Reindicating, 555, 587; summarizing table, 568
- Relations, and secondary processes, 569; example of, 422; in-, out-, intra-, extra-, and inter-, 403; one-way, 406
- Relationship vs. process, 530
- Relationships, 45; Eubank's classification of, 395; societal, 437
- Relativity, of indicators and subclasses, 145
- Reliability, 338; correlations, 912; experiment to test, 918; of S-theory, 65 ff., 158, 245
- Religious tension, 269
- Relining, 726

- Remapping, 574, 587, 727, (Eq. 69, Ch. X) 729
- Reordinating, (Eq. 46a, Ch. X) 557, 568; example of, 718; illustrated, 649; related to redispersing and reranking, 718, (Eq. 57e, Ch. X) 722
- Repopularizing, (Eq. 57c, Ch. X) 722
- Repopulating, 587, (Eq. 57a, Ch. X) 722
- Representatives, of sampling, 160
- Reranking, 566, 568, (Eq. 86, Ch. VI) 348; celeration of, (Eq. 6d, Ch. XI) 821; formula for, (Eq. 64a, Ch. X) 724; illustrated, 662
- Rescheduling, 576, 587, (Eq. 74, Ch. X) 731
- Research, fundamental problem, 919; suggestions, see Appendix III, 916
- Resisting force, 754
- Resolution, of a force, 754
- Respatiating, 587, 726
- Response, to situations, 19
- Retensing, illustrated, 625, (Eq. 57d, Ch. X) 722
- Revaluating, (Eq. 53a, Ch. X) 561; equality to monopoly, (Eq. 54c, Ch. X) 721, (Eq. 57c, Ch. X) 722; illustrated, 613, 625
- Revarying, (Eq. 58b, Ch. X) 563, 568, (Eq. 58a, Ch. X) 723; celeration of, (Eq. 6b, Ch. XI) 821
- Revoluming, illustrated, 693
- Revolution, societal, (Eq. 5, Ch. XI) 739
- Ridges, defined by equations, (Eq. 9, Ch. VII) 438; on the interrelation surface, 400
- Rules, for combining attributes, 186; for S-formulae, 67
- Rural and Urban Sociology, 456
- Rural sociology, 34
- Russian Five-Year Plan, 275
- Sampling, error, (Eq. 21a, Ch. III) 158 ff.; of S-situations, 70, 71
- Sampling error, and size of sample, (Eq. 14a, Ch. IX) 500
- Scalar product, 40, 57, 127, (Eq. 29, Ch. III) 196; formulae for, (Eq. 6, Ch. VI) 381; of two vectors, (Eq. 57, Ch. VI) 343, 911; of vectors, 309
- Scattergram, formula of, 57, (Eq. 1, Ch. VI) 381, 329
- Science, classification of, 78; content of, 4; function of, 3, 831
- Scientific method, 5; illustrative S-theory, 73; variants in, 11
- S-concepts, comprehensiveness of, 834; in understanding society, 833; objectivity of, 833; parsimony of, 834; precision of, 833
- Scriptal position, denoting an aggregation, 382
- Scripts, 32; aggregative, 46; and letters used, 192; and successive letters, 189; base for, 125; capital vs. small letters, 54, 58; pattern of, 907; person, 219; plural, (Eq. 3a, Ch. IV) 206; rules for, 904
- Secondary, 41; indicatory processes, 555; indices, 189; populations, 204, (Eq. 4c, Ch. X) 531; processes, 570, 587; temporal processes, 576
- Secondary durational processes, 730
- Secondary processes, special cases, (Eq. 54a, Ch. X) 720; terminology of, 722
- Section, of a matrix, 48
- Sector, compared with Eubank's categories, 137; definition of, 31, 58; number of, 123, 136, 185; order of presentation, 138; selection of, 450
- Sectoral S-formula, 53
- Selection, statistical, 9
- Semicolon, as operational symbol, 59, 900
- Separating, (Eq. 90a, Ch. X) 589
- Sequence of topics, 185; denotation of, 529; formula for, 706
- Sex ratio, and age, (Eq. 13b, Ch. IX) 499, 525
- S-formula, analysis illustrated, 914; for cognate social sciences, 918; descriptive vs. calculative, 105; rules for writing, 903; sectoral, 53; standardized patterns, 913; versatility of, 287
- Shortening, (Eq. 71c, Ch. X) 729
- Sigma, in S-notation, 387
- Significance, of any difference, (Eq. 17, Ch. V) 296; ratio, (Eq. 63, Ch. X) 724; ratios, (Eq. 21, Ch. III) 159
- Similarizing, dissimilarizing, and assimilizing, 544, 586
- Similizing, (Eq. 94b, Ch. X) 595
- Simple processes, 531
- Single change, definition of, 530
- Single processes, 586
- Singular, descript, 157; script, 127
- Situations, 58; defined, 903; definition of, 6; examples by institutions, 300; relativity of, 153; sample of, 73, 416; to be discovered, 921; unitary, 129; unitary, illustrated, 654, 678; unobserved but predicted, 45
- Skewing, (Eq. 66a, Ch. X) 569

- Skewness, 125, 143; discussed, 725; measure of, (Eq. 10, Ch. V) 243, 249, 586
- Slope, of the interrelational plane, 409
- S-notation, comprehensiveness of, 651; to be standardized, 922
- Social, 13; nine senses of, 24
- Social distance, critique of, 402; in Syria, 417; religions, 88, 395
- Social distance margin, 405; formula for, (Eq. 6, Ch. VIII) 437
- Social problems, 255
- Social sciences, 14
- Socializing, 586, (Eq. 93a, Ch. X) 594
- Sociating, (Eq. 6c, Ch. X) 533; limits of, 707
- Sociation, 217
- Societal, 13
- Society, 13; formula for the organization of, 435
- Sociologists, mathematical textbook for, 916
- Sociology, and statistics integrated, 848; as social science, 13; concept of, 12; corollary of the definition of, 572
- Sources, of sociological observation, 502
- Space, 26; definition of, 26, 58, 123; physical as a sociological category, 449; rule for, 903
- Space sector, notation of, 451; see Part IV, 447
- Spacing, (Eq. 3c, Ch. X) 531, 573, 586
- Sparsing, 729
- Spatial, case script, 52; class-interval, 50; correlation, (Eq. 41, Ch. VI) 335; densities, (Eq. 13, Ch. VIII) 491; distributing, 573; durations, (Eq. 17, Ch. IX) 504; interrelations, 397, 726; terminology of process, 728
- Spatial dimensioning, four classes, (Eq. 67, Ch. X) 725, 726
- Spatial exponents, (Eq. 2, Ch. II) 124; rules for, 905
- Spatial processes, 573; conspectus of, 726; double classification of, 573; subclassified, 728
- Spatializing, 586, 726
- Spatiating, 586, 726
- Special case, S-theory, 59
- Specialists, as defining plurals, 214
- Specifications, of this volume, 24
- Specificity, (Eq. 45, Ch. VI) 338
- Sporting, 587, (Eq. 101, Ch. X) 734
- S-situation, 31; sample of, 167, 224, 279; see 300 illustrations appended to Chs. II-XII
- S-space, and subspaces, 68
- $\Sigma$  symbol, 510
- S-system, in understanding society, 839
- Standard deviations, 161, 243; in monopoly, (Eq. 19a, Ch. V) 297, 899
- Standard error, for a force and its factors, (Eq. 16, Ch. XI) 823; for societal force, 760; formula for, (Eq. 14b, Ch. IX) 525; momentum, (Eq. 88, Ch. X) 732; of a standard deviation, (Eq. 16b, Ch. V) 296; of arithmetic mean, (Eq. 13, Ch. V) 295; of difference, (Eq. 18, Ch. V) 296; of Kurtosis, (Eq. 16d, Ch. V) 296; of skewness, (Eq. 16c, Ch. V) 296; to be derived, 922
- Static situations, 36
- Statistical dimensionality, of space, 489
- Statistical series, 185
- Statistics, 19; and sociology integrated, 848
- Status relationship, 396
- Step-interval, 15
- S-theory, a mnemonic phrase for, 128; adequacy of, 70; algebraic statement, 59; and psychology, 640; and understanding phenomena, 832; appraisal of, see Ch. XII, 831; characterized, 8, 14; classification in, 62; contribution toward controlling phenomena, 854; criticism of, 918; development of, 120; early form of, 130; formulae, 73; fruitfulness of, 63; geometry of, 67, 120; glossary of symbols, 899; in all social sciences, 392; limitations of, 842; notation as methodological hypotheses, 463; parsimony of, 60; precision of, 61, 128; preliminary statement of, 25; reliability of, 65 ft., 128; revision of, 130; utility of, 60, 917; verbally stated, 58; verification of, 64
- Straight line, equation of, (Eq. 12, Ch. VI) 319
- Stratification, 249
- Stratifying, 587, (Eq. 100a, Ch. X) 734
- Structure, of a "loose" situation, 519
- Subclasses, 145
- Subclassification, 86, 207
- Subclassifying sign, 54
- Subindicators, 29
- Subnormality, methods of reducing, 258
- Subnormals, (Eq. 26b, Ch. V) 254
- Subtracting, 26, 59; of attributes, (Eq. 5, Ch. III) 186
- Summary, of S-theory, 58
- Summative combination, 732

- Summum bonum, 831  
 Superordination, 249  
 Supernormals, (Eq. 26d, Ch. V) 254  
 Surfaces, biplanar, 404; interrelational, 398; interrelational formula, (Eq. 17, Ch. VII) 442; goodness-of-fit of, 412; monoplanar, 408; polyplanar, 411  
 Symbolic substitution, 857  
 Symbols, and data, 598, 893; clarifying of, 23; glossary of, 899; need for, 20, 21; number of in S-theory, 18; of S-theory, 31; precision of, 24; resistance to, 22; the difficulty of learning, 20; usual mathematics, 127; verbal vs. algebraic, 19  
 Synthesis, of sectors, see Part VI, 829  
 System, 8; concept of, 18  
 Systematics, of the quantic formula, (Eq. 11, Ch. VI) 315  
 Systematization, 6; research upon, 916  
 T, see Time, 31  
 Temporal, case scripts, 51; class-intervals, 49; class script, 48; densities, 463; exponent, 495; exponent, rules for, 904; index, secondary, (Eq. 7, Ch. IX) 524; interrelations, 397; scripts, (Eq. 2.3, Ch. IX) 522  
 Tensing, (Eq. 35, Ch. X) 550, 587, 714  
 Tension, and cause and force, 759; and societal processes, 276; hypothesis, (Eq. 39a, Ch. X) 554, 921; index, (Eq. 34a, Ch. V) 265; measured, 653; numerical examples, 266  
 Tension equation, an illustrative situation, 802  
 Tension theory, 108, 121, 262; explored, 920; illustrated, 779, 802, 820; indicant in, (Eq. 37, Ch. V) 304; lack of concepts, 305; matrix equation of, (Eq. 49, Ch. V) 273; simple, (Eq. 34b, Ch. V) 265  
 Tensions, illustrating, correlation of, 635  
 Tensor theory, 918  
 Terminology, for processes, 529; of spatial processes, 728  
 Ternary, celerations, 741; process, 531; spatial processes, 726  
 Tertiary, 41, (Eq. 4d, Ch. X) 531; indicative processes, 568; processes, 571, 587  
 Tetrads, 326  
 Theory, 8; in Sociology, 121  
 Therapy, social, 257  
 Threatening, 592  
 Time, 26; conception of, 524; curves, 914; definition of, 27, 123; rate, 34; Sector, see Part V, 493; series, 56; series, correlation of, (Eq. 89, Ch. VI) 349; symbol for, 31; units of, 48, 58  
 Timeless situations, 36  
 Timing, (Eq. 3d, Ch. X) 531, 574, 586; celeration of, 741  
 Tolerating, 586, (Eq. 92b, Ch. X) 593  
 Tools, 149  
 Topographical, index, 463; indicators, 191  
 Topology, 129; illustration of, 673  
 Transformations, from intercorrelations, 353; utility of, 320  
 Transportation systems, 149  
 Trend, 349  
 Trending celeration, (Eq. 2, Ch. XI) 737  
 Triangular matrix, 404  
 Truth, two tests of, 64  
 Type, symbolized, 104, 430  
 Typical persons, reduction to, 323  
 Underlining, 900; denoting indefiniteness, 77, 912  
 Understanding, as function of science, 3; of phenomena, 832  
 Unemployment, formulae relating, 78  
 Uniformity of nature, assumption of, 383  
 Unifying, 586, (Eq. 90b, Ch. X) 589  
 Units, common and diverse, 190; comparable, 161; comparable for P, 750; converted from ordinal to cardinal, 490; development of, 161; of ethical evaluation, 183; of force, 745, 746; of observation, 32, 122; of tension, 305; of value, 91; of varying precision, (Eq. 16, Ch. III) 153; situation hypothesis, 129; specified by class-interval, 50; symbols for, 126  
 Univalent relation, 406  
 Unknown situations, predicted, 45  
 Unknown, in equations, 59  
 Unlimited, desideratum, (Eq. 47b, Ch. V) 272  
 Unreliability, of analysis, 171  
 Urban, sociology, 34  
 Utility, of S-theory, 60; of S-theory to be explored, 917  
 Validity, 19, 160; research upon, 918  
 Valuating, (Eq. 30, Ch. X) 547; illustrated, 613  
 Value, attributes, (Eq. 25a, Ch. V) 253; index, 899; indicants, (Eq. 25b, Ch. V) 253; of indicators, 191; plurals, 218; see Desiderata, 29

- Variance, 143, (Eq. 9, Ch. V) 243; of interrelations, 392, 587; of true score, (Eq. 47, Ch. VI) 338; of variables, (Eq. 120, Ch. VI) 357
- Varying, 587, (Eq. 94a, Ch. X) 595
- Variation, (Eq. 21c, Ch. V) 249
- Vector, addition, subtraction, and multiplication of, 195; and operators, 164; and scripts, 162 ff.; defined, 67 ff.; division of, 198; examples of transformation, 384 f.; formula, 164; oblique and orthogonal, 346; of S-theory, 119; sheaf or ray, 163; transformation of, 319
- Vector product, of two vectors, (Eq. 67, Ch. IV) 344, 918
- Vectorial formula of a distribution, (Eq. 5, Ch. V) 306
- Velocity, 35; during, 76; of change, (Eq. 3a, Ch. IX) 495; standard error of, (Eq. 16, Ch. XI) 823
- Verification, 8, 19; of a classification by experiment, 394; of S-theory, 64; research suggestions, 918
- Vital statistics, equation of, (Eq. 11, Ch. X) 536; valid principles, 537
- Volumating, illustrated, 693, 726
- Voluming, (Eq. 67d, Ch. X) 725, 726
- Volumizing, 726
- Wants, see Desiderata, 30
- Weapons, 149
- Wishes, see Desiderata, 30
- Zero, absolute, 155; origin, 176; person script, 37; point, for a force, 744









Not to be issued

Acc. No.	19381
Class No.	D. 1.
Book No.	43

RARE BOOK  
NOT TO BE ISSUED

